University College London

Doctoral Thesis

# The e-MERLIN L-Band Legacy Survey of Cygnus OB2

*Author:*
Jack Morford

*Supervisor:*
Prof. Raman Prinja

*A thesis submitted in fulfilment of the requirements*
*for the degree of Doctor of Philosophy*

*in the*

Astrophysics
Department of Physics & Astronomy

March 20, 2017

*For my Grandparents, George and Iris Hart...*

# Declaration of Authorship

I, Jack Morford, declare that this thesis titled, "The e-MERLIN L-Band Legacy Survey of Cygnus OB2" and the work presented in it are my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

"*We are like butterflies who flutter for a day and think it is forever*"

Carl Sagan

# *Abstract*

The Cygnus OB2 Radio Survey (COBRaS) is an e-MERLIN Legacy survey awarded $\sim$ 300 hours worth of observing time. Split between L-(42 hrs) and C-(252 hrs) band, the project aims to push the limits of e-MERLIN's enhanced capabilities by providing highly sensitive, targeted, deep-field radio maps of the core of the Cygnus OB2 association. The L-band (21cm) observations, completed throughout the Spring of 2014, serve as the focus of the work presented here.

This thesis primarily aims to investigate the radio emission from massive ($M \gtrsim 8$ M$_\odot$) stars in the Cyg OB association, in order to obtain a better understanding of their mass-loss rates and binarity. A description of the entire data reduction process of the COBRaS 21cm observations is given, as is a discussion of the mitigation of radio frequency interference (RFI). In the advent of new radio telescopes, the technical skills, research and tools developed throughout this process aims to serve the wider radio astronomy community.

Further technical work is presented in the form of novel source extraction and cross-correlation scripts that deliver the COBRaS L-band all source catalogue (CLASC). Reaching resolutions of $\sim$180 mas, precise positions and 21cm flux densities are presented for 41 sources found within a 0.48 deg$^2$ region, centred on the Cyg OB2 association. In obtaining a $3\sigma$ sensitivity limit of $\sim$ 100 $\mu$Jy, these observations significantly improve upon previous surveys of the region. Accurate determinations of the mass-loss rates of single massive OB stars are used to investigate their stellar wind properties, helping to resolve current discrepancies in the mass-loss rates via clumped and structured hot star winds. Finally, an analysis of the non-thermal radiation from the colliding winds of massive star binaries seeks to clarify our understanding of such systems and their radio emission processes.

# *Acknowledgements*

First and foremost, I would like to express my thanks to Danielle, whose patience, understanding and attention to detail has guided me through much of this PhD. She has been approachable, caring and has had an answer to every question I have thrown at her. I cannot thank her enough for the countless discussions and interesting debates on all matters of radio astronomy and more. Secondly, a huge thanks goes to Raman, whose passion for scientific endeavour and knowledge of astrophysics has never ceased to amaze me. He is both an inspiration and a role model, and I am very grateful for his guidance throughout my PhD. Many thanks also go to Jeremy, for not only introducing me to the wonders of radio interferometry but for keeping a watchful eye over my progress through this PhD. Furthermore, I would like to add thanks to Luke, who made my transition from a masters to postgraduate student incredibly comfortable, both scientifically and socially.

In a wider context, I'd like to thank my parents, all four of which have given me the love, support and encouragement that has not only made much of this thesis possible, but made me the person I am today. It is difficult to express how much I appreciate your belief in me, especially in times of doubt. I am extraordinarily lucky and extremely proud to have you all, I cannot thank each of you enough. In addition, thanks go to my close friends, each of whom have supported me throughout some stage of my PhD. To my fellow musicians and bandmates, for allowing me to blow off some steam in times of need. To Greig, who always managed to knock me back down to Earth; and to my brothers and sisters, for making me feel part of something. A special thank you goes to Laura, for the unquestionable love and support she has given me, through the very hard times and the very good. Without her hard work and caring attitude, this document would not be in the state it is in today. She has been my rock, and I only hope to return the favour in the near future.

Finally, I would like to express thanks to my Grandfather, who sadly passed away throughout the time of writing this thesis. He has been a huge influence on my life and had warmest heart of any man I have ever known. He will be sorely missed.

# Contents

**Declaration of Authorship** **5**

**Abstract** **8**

**Acknowledgements** **9**

**1 Introduction** **23**

1.1 Massive stars . . . 24

1.1.1 Basic properties . . . 25

1.1.2 The formation of massive stars . . . 28

1.1.3 The evolution of massive stars . . . 29

1.1.4 The role of massive stars . . . 32

1.2 Massive star winds . . . 33

1.2.1 Radiative line-driven wind theory . . . 37

1.2.2 Structure in the winds of hot stars . . . 45

1.2.2.1 Small-scale structure - clumping . . . 46

1.2.2.2 Co-rotating interaction regions - CIRs . . . 50

1.3 Massive stellar clusters . . . 53

1.3.1 Fundamental properties of massive stellar clusters . . . 54

1.3.2 The formation, evolution and survival of MSC . . . 57

1.3.2.1 Formation of young stellar systems . . . 57

1.3.2.2 Evolution and survival of young stellar systems . . . 59

1.4 Cygnus OB2: a massive OB association . . . 61

1.4.1 Location . . . 62

1.4.2 Visual extinction in Cyg OB2 . . . 65

1.4.3 Distance to Cyg OB2 . . . 66
1.4.4 Cluster members . . . 68
1.4.5 General properties . . . 70
1.5 The e-MERLIN array . . . 73
1.6 COBRaS - an e-MERLIN Legacy survey . . . 74
1.6.1 Scientific aims . . . 75
1.6.1.1 Mass-loss, stellar winds, and massive star evolution . . . 75
1.6.1.2 Binarity and the incidence of non-thermal radiation. . . . 77
1.6.2 Technical aspects . . . 79
1.6.2.1 Observation strategy . . . 79
1.6.2.2 Data treatment . . . 80
1.7 The aims and structure of this thesis . . . 82

**2 e-MERLIN COBRaS L-band data treatment and reduction 85**
2.1 An introduction to radio interferometric theory . . . 85
2.1.1 The two element interferometer . . . 87
2.1.2 A realistic radio interferometer . . . 91
2.2 Radio frequency interference . . . 95
2.2.1 SPPlot . . . 96
2.2.2 SERPent . . . 101
2.2.3 RFI removal process for COBRaS L-band data set . . . 103
2.3 Calibration of COBRaS L-band data set . . . 113
2.3.1 Calibration outline . . . 113
2.3.2 Delay correction and flux density scale . . . 116
2.3.3 Phase calibration . . . 120
2.3.4 Amplitude calibration . . . 121
2.3.5 Bandpass calibration . . . 124
2.3.6 Self-calibration . . . 126
2.4 Imaging of the COBRaS Legacy L-band data . . . 133
2.4.1 Imaging procedure and the clean algorithm . . . 133
2.4.2 Initial images and artefact hunting . . . 136

2.4.3 Production of the COBRaS wide-field L-band maps . . . . . 141
2.5 Conclusions and future work . . . . . 144
2.5.1 A summary of the L-band data reduction process . . . . . 144
2.5.2 Assessment of the data calibration process . . . . . 145
2.5.3 Future work . . . . . 146

**3 The COBRaS L-band all source catalogue 151**
3.1 Source detection and flux extraction . . . . . 151
3.1.1 Source Extraction Algorithm for COBRaS - SEAC . . . . . 153
3.1.1.1 The floodfill algorithm . . . . . 155
3.1.1.2 Background noise map . . . . . 156
3.1.1.3 Source position determination . . . . . 158
3.1.1.4 Source flux determination . . . . . 159
3.1.1.5 SEAC outputs . . . . . 160
3.1.2 Search and destroy - SAD . . . . . 161
3.2 Catalogues and cross-correlation techniques . . . . . 166
3.2.1 The Cyg OB2 super catalogue . . . . . 166
3.2.1.1 Cyg OB2 OB star catalogue . . . . . 170
3.2.1.2 Cyg OB2 OB candidate catalogue . . . . . 170
3.2.2 Other catalogues considered for cross-correlation . . . . . 172
3.2.2.1 2MASS catalogue . . . . . 172
3.2.2.2 Optical photometry GTC/OSIRIS catalogue . . . . . 174
3.2.3 Cross correlation script . . . . . 174
3.3 COBRaS L-band source population catalogue . . . . . 176
3.3.1 Source detection . . . . . 176
3.3.2 Source identification . . . . . 197
3.3.3 Source counts . . . . . 200
3.4 Analysis of the COBRaS 21cm detections . . . . . 202
3.4.1 Identified objects in CLASC . . . . . 202
3.4.1.1 CLASC ID: #2 . . . . . 204
3.4.1.2 CLASC ID: #7 . . . . . 204

3.4.1.3 CLASC ID: #9 . . . 206
3.4.1.4 CLASC ID: #10 . . . 206
3.4.1.5 CLASC ID: #11 . . . 207
3.4.1.6 CLASC ID: #17 . . . 208
3.4.1.7 CLASC ID: #18 . . . 209
3.4.1.8 CLASC ID: #27 . . . 209
3.4.1.9 CLASC ID: #33 . . . 209
3.4.1.10 CLASC ID: #34 . . . 210
3.4.1.11 CLASC ID: #35 . . . 211
3.4.1.12 CLASC ID: #38 . . . 211
3.4.1.13 CLASC ID: #39 . . . 212
3.4.1.14 CLASC ID: #41 . . . 212
3.4.2 Un-identified objects in CLASC . . . 213
3.4.3 Discussion . . . 214
3.5 Conclusions and future work . . . 217

**4 21cm (L-band) observations of single massive stars in Cyg OB2 221**
4.1 The winds of hot stars: a summary . . . 222
4.1.1 The importance of stellar mass-loss and its current uncertainty . . . 223
4.2 Radio emission from massive stellar winds . . . 226
4.2.1 The winds of massive stars . . . 226
4.2.2 The thermal free-free emission spectrum from early-type stars . . . 227
4.2.3 The effect of clumping on the thermal free-free emission spectrum . . . 230
4.2.4 Predicting the radio flux from massive star winds . . . 231
4.3 Mass-loss rates from single massive stars in Cyg OB2 . . . 235
4.3.1 Single massive star sample selection . . . 235
4.3.2 Adopted fundamental parameters . . . 236
4.3.3 Cygnus OB2 #7 . . . 239
4.3.4 Cygnus OB2 #8C . . . 241
4.3.5 Cygnus OB2 #12 . . . 242
4.3.6 The remaining sample selection . . . 246

4.4 Discussion and conclusions . . . 247
4.4.1 A look ahead to the COBRaS C-band data . . . 249

**5 21cm radio emission from massive star binaries in Cygnus OB2 253**

5.1 Massive star binaries . . . 253
5.1.1 The impact of massive star binaries . . . 256
5.2 Non-thermal radio emission from massive star binaries . . . 258
5.2.1 Synchrotron radiation . . . 259
5.2.2 Colliding wind regions . . . 261
5.3 Massive star binaries within the COBRaS 21cm radio observations . . . 264
5.3.1 Target sample selection . . . 264
5.3.2 MT267/A11 (ID: #8) . . . 269
5.3.3 Cyg OB2 #8A (ID: #22) . . . 270
5.3.4 Cyg OB2 #9 (ID: #21) . . . 276
5.3.5 Cyg OB2 #5 (ID: #6) . . . 280
5.3.5.1 The NE component of Cyg OB2 #5 . . . 284
5.3.5.2 The SW component of Cyg OB2 #5 . . . 286
5.3.6 The undetected massive star binary selection . . . 291
5.3.7 Unidentified potential PACWB systems . . . 291
5.4 Discussion and conclusions . . . 295

**6 Summary, conclusions and future work 301**

6.1 Thesis summary . . . 301
6.2 The next steps in the COBRaS project . . . 304
6.2.1 L-band data . . . 304
6.2.2 C-band data . . . 306
6.3 Future proposals . . . 308

**A AIPS functions and nomenclature 311**

**B The $u, v$ coverages of the seven COBRaS 21cm pointings. 315**

**C Additional information regarding the CLASC 319**

**Bibliography** **321**

# List of Figures

1.1 A Hertzsprung-Russell diagram showing the spectral classification of stars as a function of luminosity and temperature . . . . . 27
1.2 The evolutionary tracks of massive stars (12-120 $M_{\odot}$) . . . . . 31
1.3 An example of a P-Cygni profile . . . . . 34
1.4 Plots to highlight the results from 1-D (left) and 2-D (right) smooth source function (SSF) models of the line-deshadowing instability in hot stars . . . . 44
1.5 A model of the radial stratification of the clumping factor . . . . . 48
1.6 A comparison of the derived $\dot{M}$ values from $\rho^2$ diagnostics with those derived using the Pv resonance line for a sample of 40 Galactic O-type stars . . . . 49
1.7 A dynamic spectrum of the Si v $\lambda$1393, 1402 resonance doublet for the O4If(n) star $\zeta$ Pup as observed as part of the IAU MEGA campaign . . . . 51
1.8 A 2-D greyscale plot of a CIR model . . . . . 53
1.9 Relative positions of the 9 OB associations and 10 young open clusters in the Cygnus region . . . . . 63
1.10 An $H_{\alpha}$ image of the Cygnus-X region . . . . . 64
1.11 A visual extinction map in galactic coordinates of the Cygnus X region . . . 67
1.12 Results of the high-precision proper motion (PM) study of 798 stellar members of Cyg OB2 . . . . . 71
1.13 The positions of the seven antennas across the UK that make up the e-MERLIN radio interferometer . . . . . 73
1.14 Positions for the (a) seven L-band (21cm) and (b) 42 C-band (6cm) COBRaS pointings . . . . . 81

2.1 A cartoon depicting the response or primary beam pattern of a given antenna 86
2.2 A schematic illustration of a simplified two element interferometer . . . . . 88

2.3 A schematic illustration of a more realistic two element interferometer . . . 92
2.4 To highlight the differences in the aesthetics of SPFLG and SPPlot . . . . . 97
2.5 A typical output from SPPlot . . . . . 98
2.6 A typical output from SPPlot in its combined format . . . . . 100
2.7 Plots highlighting the application of the e-MERLIN flagmask . . . . . 106
2.8 An amplitude plot of the RR polarisation, 1-5 (Lo-Kn) baseline of the phase-calibrator source . . . . . 108
2.9 A clean amplitude plot of the RR polarisation, 1-5 (Lo-Kn) baseline of the phase-calibrator source . . . . . 110
2.10 A flow diagram describing the steps taken in the calibration of the COBRaS Legacy L-band data . . . . . 115
2.11 Amplitude and phase plots created using the AIPS task POSSM, pre- and post delay calibration . . . . . 119
2.12 POSSM and VPLOTs post phase calibration . . . . . 122
2.13 A POSSM plot of the RR polarisation of the 1-5 baseline of the phase calibrator post A&P calibration . . . . . 125
2.14 Two POSSM plots to show the effect of the bandpass calibration . . . . . 127
2.15 A flow diagram highlighting the steps taken throughout self-calibration . . . 129
2.16 An initial image of the phase-calibrator, J2007+404 before any self-calibration attempts . . . . . 130
2.17 A comparison between two images of the phase calibrator, pre- and post self-calibration . . . . . 132
2.18 Successive images produced from the self-calibration routine of the target field pointing B . . . . . 134
2.19 Three images of the phase calibrator to highlight various image artefacts . . 137
2.20 Initial wide field images of the target field pointing E . . . . . 139
2.21 Two images of SHBW90 taken from pointing D showing image artefacts around the source . . . . . 140

3.1 The calculated noise map created within SEAC in units of 5 × the local RMS 157
3.2 A typical example of the output plots from SEAC . . . . . 162

3.3 Four examples of different sources within the COBRaS L-band field of view 163
3.4 The positions of the 14355 sources within the Cyg OB2 Super Catalogue . . 169
3.5 (J-H, H-K) diagram of the sources from the Cyg OB2 super catalogue . . . 171
3.6 (J-H, H-K) diagram of the sources from the 2MASS catalogue . . . . . 173
3.7 The positions of the 64157 sources from the GTC/OSIRIS observations from Guarcello et al. (2012) . . . . . 175
3.8 The regions covered by the seven wide field COBRaS L-band images on the sky . . . . . 177
3.9 The positions of the 41 sources (shown in red) derived from the CLASC . . 184
3.10 CLASC sources detected in both epochs . . . . . 185
3.10 CLASC sources detected in both epochs . . . . . 186
3.10 CLASC sources detected in both epochs . . . . . 187
3.10 CLASC sources detected in both epochs . . . . . 188
3.10 CLASC sources detected in both epochs . . . . . 189
3.10 CLASC sources detected in both epochs . . . . . 190
3.10 CLASC sources detected in both epochs . . . . . 191
3.10 CLASC sources detected in both epochs . . . . . 192
3.10 CLASC sources detected in both epochs . . . . . 193
3.11 CLASC sources only detected in the Apr. 26th observations . . . . . 194
3.11 CLASC sources only detected in the Apr. 26th observations . . . . . 195
3.12 CLASC sources only detected in the Apr. 26th observations . . . . . 196
3.13 Normalised differential source counts plot . . . . . 203

4.1 Fluxes obtained for the B0 Ia star $\epsilon$ Ori from observations in optical to the radio wavelengths . . . . . 224
4.2 A schematic diagram to illustrate the regions in the wind which are dominated by emission at a given waveband . . . . . 226
4.3 A schematic illustration to highlight the geometry of the model used to derive the free-free thermal emission spectrum throughout the winds of early-type stars . . . . . 228

4.4 Theoretical predictions of mass-loss and 21cm radio emission from massive stars as a function of spectral-type . . . 233
4.5 A blank field image of Cyg OB2 #7 . . . 239
4.6 A blank field image of Cyg OB2 #8C . . . 242
4.7 Two images of the hypergiant Cyg OB2 #12 from the COBRaS 21cm observations . . . 243
4.8 A semi-empirical model of the radial stratification of $f_{cl}$ as a function of the wavelength of free-free emission within the winds of hot stars . . . 250

5.1 Relative fractions of various binary interaction scenarios . . . 257
5.2 Typical synchrotron emission spectrum . . . 260
5.3 Wind collision region schematic . . . 261
5.4 Images of A11 from the COBRaS observations . . . 270
5.5 21cm radio light curve of the PACWB A11 . . . 271
5.6 Images of Cyg OB2 #8A from the COBRaS observations . . . 272
5.7 21cm radio light curve of the PACWB Cyg OB2 #8A . . . 274
5.8 21cm and 6cm radio light curve of the PACWB Cyg OB2 #8A . . . 275
5.9 Images of Cyg OB2 #9 from the COBRaS observations . . . 277
5.10 21cm radio light curve of the PACWB Cyg OB2 #9 . . . 277
5.11 21cm and 6cm radio light curve of the PACWB Cyg OB2 #9 . . . 279
5.12 Cartoon depicting the different components of Cyg OB2 #5 . . . 282
5.13 Images of Cyg OB2 #5 from the COBRaS observations . . . 282
5.14 Images of Cyg OB2 #5 with a $u, v$ taper applied, from various observations 285
5.15 21cm, 6cm and 3.6cm radio light curve of the PACWB Cyg OB2 #5 . . . 289

B.1 $u, v$ coverages of pointings A-C from the Apr. 26th and Apr. 11th epochs . 316
B.1 $u, v$ coverages of pointings D-F from the Apr. 26th and Apr. 11th epochs . 317
B.1 $u, v$ coverages of pointings G from the Apr. 26th and Apr. 11th epochs . . . 318

# List of Tables

1.1 A comparison between the fundamental parameters of the three outlined stellar clusters . . . 55
1.2 To highlight the main observations and surveys of Cyg OB2 . . . 69
1.3 The fundamental physical parameters that describe the Cygnus OB2 association . . . 70
1.4 The basic observing capabilities of the e-MERLIN array . . . 74

2.1 Basic information regarding the COBRaS L-band Legacy data set . . . 104
2.2 SERPent flagging parameters used in the treatment of RFI . . . 107
2.3 The observed sources, their data sizes and integration times across both COBRaS observation epochs . . . 112
2.4 The antenna elements of the e-MERLIN array . . . 114
2.5 The known flux at the central frequency of each of the 8 L-band IFs of the flux-calibrator source 1331+305 . . . 117
2.6 The derived fluxes of the point and phase calibration sources of the COBRaS L-band Legacy data set . . . 123
2.7 The final statistics of the fully calibrated COBRaS 21cm datasets . . . 147

3.1 To highlight the differences between the positions and integrated fluxes as obtained via SEAC and SAD . . . 164
3.2 A list of the previous catalogues containing information on the members of Cyg OB2 . . . 168
3.3 The COBRaS L-band all source catalogue . . . 183
3.4 Information regarding the specific catalogues from which sources from the CLASC have been identified with . . . 199

3.5 Source identification information . . . . . 201
3.6 Source counts from the CLASC . . . . . 202
3.7 CLASC sources without identification . . . . . 214

4.1 COBRaS L-band measured flux densities for the sample of nine OB stars in Cyg OB2 . . . . . 237

5.1 Positional and orbital information for 27 PACWB systems in the field of view of the COBRaS 21cm observations . . . . . 266
5.2 Positions and fluxes of four PACWB systems detected in the COBRaS 21cm observations . . . . . 267
5.3 Phase information of the four multiple star systems observed within the COBRaS observations . . . . . 268
5.4 Phase information of the SW component of Cyg OB2 #5 from four different models . . . . . 287
5.5 Information regarding the eight potential PACWB systems . . . . . 294

A.1 The definitions of the shorthand AIPS table/file names . . . . . 311

C.1 Positional information and their associated uncertainties of the 41 objects within CLASC . . . . . 320

# Chapter 1

# Introduction

This thesis is based on research carried out primarily on the data from Legacy L-band (21cm) observations from COBRaS: The Cyg OB2 Radio Survey. The UCL led, e-MERLIN (extended Multi-Element Radio Linked Interferometer Network) Legacy project (P.I. Prof Raman Prinja) aims to compile the most substantial radio dataset to date of the most massive OB association in the northern hemisphere, Cygnus OB2. In providing highly-sensitive, targeted deep-field maps of the exceedingly OB rich association, COBRaS aims to better our understanding of massive stars, massive stellar clusters (MSC) and the other stellar content that exists within them. This thesis provides the steps taken in order to achieve these aims, outlining the scientific motivation behind such an ambitious project, developing the techniques used, and detailing the analysis from which significant results are both discussed and concluded upon. In addition to addressing major scientific uncertainties in the study of massive stars, the technical work presented below will benefit not only the future work carried out within COBRaS but also other e-MERLIN groups in the wider community. In the advent of the Square Kilometre Array (SKA), radio astronomy is entering into a new era and projects such as COBRaS and the work presented here are of upmost importance, not only to highlight the incredible science one can achieve at radio wavelengths, but to learn, develop and infer new techniques to cope with extremely large datasets.

This introduction will proceed by first exploring the 'bigger' picture in relation to this thesis with a brief introduction to massive stars including their basic properties, evolution and their effect on the surrounding environment. In particular some major attention is given

to the winds of massive stars throughout Section 1.2. After which some focus is given to the impact of massive stars within massive stellar clusters (Section 1.3) by drawing attention to the basic properties of such systems including their formation and evolution. This will be followed by a more detailed discussion of Cyg OB2 itself as a massive OB association (Section 1.4). A description of the e-MERLIN array will follow (Section 1.5), proceeded by a general overview of the COBRaS project (Section 1.6). Finally, the chapter is concluded with a more detailed discussion upon the contents, aims and objectives of this thesis (Section 1.7).

## 1.1 Massive stars

A star can be considered a luminous sphere of plasma in hydrostatic equilibrium, meaning the inward pull of gravity is balanced by the outward radiation and thermal pressure forces. Their source of energy is a result of thermonuclear processes within their interiors, whereby lighter elements are fused together to form heavier elements. A bi-product of this process is the production of energy, which propagates either convectively, or radiatively outwards through the photosphere of the star (i.e. the stellar surface). This energy is radiated away in the form of electromagnetic radiation, i.e. photons with a large range of frequencies. Their most important parameter is undoubtedly their mass, the extent of which can affect their temperature and hence the various processes leading to their energy production. Changes to their energy production processes will alter their physical properties and hence their subsequent evolution (and therefore lifetime) and can have a dramatic consequence on their eventual end point.

Massive stars include all stellar objects above a mass of $\sim 8$ $M_{\odot}$. Whilst a number of these objects have been observed up to a mass of $\sim 150$ $M_{\odot}$, the object RMC 136a1 is currently the most massive known star observed in the universe, with a derived mass of $\sim 315$ $M_{\odot}$ (Crowther et al., 2016). A fundamental difference between these groups of stars in comparison to their lower mass ($M \lesssim 8$ $M_{\odot}$) counterparts is their eventual end point. A low mass stellar object, having slowly exhausted the hydrogen (H) and helium (He) within their cores, will slowly fade away becoming an electron degenerate white dwarf. Massive

stars on the other hand, able to reach much higher temperatures within their cores, can form a degenerate iron core, the collapse of which leads to the production of a proto-neutron star. The subsequent evolution of this object can be hugely dramatic and can result in a number of different scenarios including a (core-collapse) supernova (SN), leaving behind a neutron star remnant. The core can also undergo a direct collapse to black hole or result in the production of a long duration gamma-ray burst (GRB; Crowther 2012).

### 1.1.1 Basic properties

In order to categorise the vast range of stellar object's observed within our universe, a given star can be classified according to its stellar spectrum. Observations of a star's spectrum (i.e. a measure of the radiation received as a function of wavelength) will reveal the relative strength of the different emission (and absorption) features and thus indicate the relative abundance of an atom or ion. Since such ion abundances are dictated by the density and temperature of the emitting regions (i.e. the photospheric or effective temperature of the star, $T_{eff}$), certain characteristics of a stars spectrum can be used to give an objective measure of its $T_{eff}$ and density. The common spectral classification scheme is known as the Morgan-Keenan (MK) system and was developed by Morgan, Keenan, and Kellman (1943). This system was based upon the older Harvard system whereby each star is assigned a spectral classification (in the form of a single letter from the alphabet) and a luminosity class (using Roman numerals). The spectral (alphabetic) classification takes the form of $OBAFGKM$ in order from the hottest (O-type) to the coolest (M-type). In the modern MK system, each letter is further sub-divided into numeric divisions from 0-9, with 0 denoting the hottest for that given subclass. The proceeding Roman numerals describe the luminosity class of a given star which provides a measure of the objects surface gravity (and hence density). The spectral lines of dense stars with a higher surface gravity will be subject to a greater amount of pressure broadening. A measure of this broadening leads to additional luminosity classes, which are distinguished as follows:

- 0 or Ia$^+$: extremely luminous supergiants or hypergiants
- Ia: luminous supergiants

- Iab: intermediate luminous supergiants
- Ib: less luminous supergiants
- II: bright giants
- III: normal giants
- IV: subgiants
- V: main-sequence stars (dwarfs)
- sd: sub-dwarfs.
- D: white dwarfs.

The MK spectral classification system of stars characterises these objects according to their temperature and luminosity. These classifications can be plotted as a function of temperature and luminosity on a Hertzsprung-Russell diagram (HRD) as shown in Figure 1.1. In regards to massive stars (i.e. all those above a mass $M \gtrsim 8$ $M_{\odot}$), these only encompass those objects with either an O- or early B-type spectral classification. Early B-type stars range in mass from $\sim 8-20$ $M_{\odot}$ and O-type stars are those with a mass larger than $\sim$20 $M_{\odot}$. They contribute to only $\sim 1\%$ of the entire stellar content (by mass) of a new born population of stars within a galaxy (Crowther, 2012). However, due to their high bolometric luminosities ($\gtrsim 10^4 - 10^6 L_{\odot}$), they dominate the appearance of star-forming galaxies whilst their ionising UV radiation has a profound impact on their surrounding environment, leading to the production of HII regions (i.e. regions of ionised inter-stellar hydrogen).

Massive stars are relatively short lived in comparison to their less-massive counterparts. The lifetime of a given star is proportional to its mass divided by its luminosity, i.e. its lifetime is dependant on the rate of energy radiated (i.e. its luminosity) over how much energy it physically has to radiate (i.e. its mass). This scales with the inverse square of its stellar mass (if assuming a mass-luminosity relation $L \propto M^{\alpha}$ where $\alpha \approx 3$; Crowther 2012), meaning massive stars radiate a larger amount of their total energy much faster than 'normal' (less-massive stars). Typical lifetimes for massive stars are a few tens of

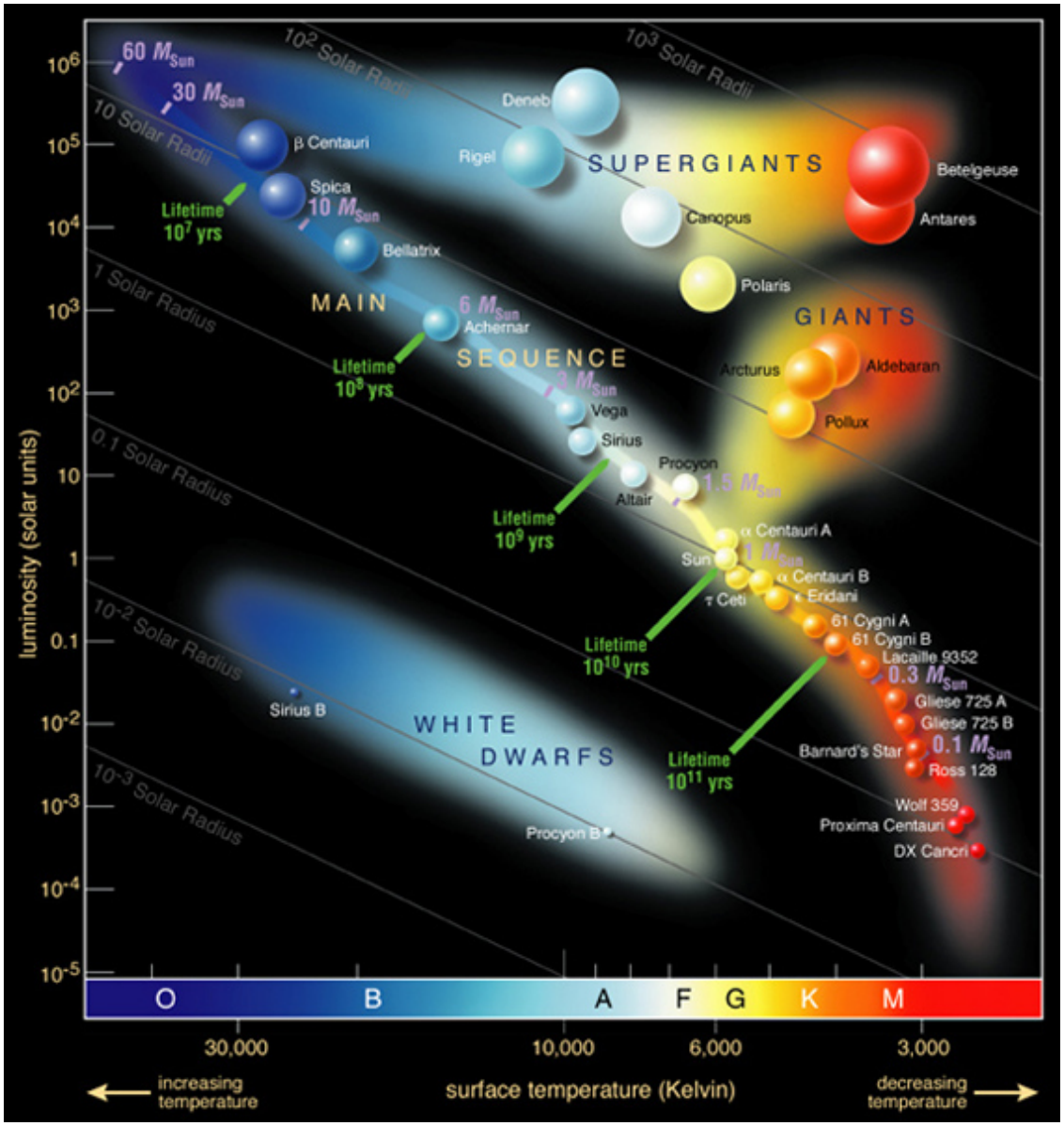



FIGURE 1.1: A Hertzsprung-Russell diagram showing the spectral classification of stars as a function of luminosity and temperature. The main sequence (MS), the phase in which stars spend the majority of their lives, is highlighted by the large diagonal strip across the image (from the top left to the bottom right) Image credit: European Southern Observatory (ESO), *url*: https://www.eso.org/public/images/eso0728c/.

millions of years. In comparison, a G-type star ($\sim 1 M_{\odot}$) such as our star is expected to live for $\sim 10$ billion years.

### 1.1.2 The formation of massive stars

The majority of stars have been observed to reside (and hence were likely born) within clustered environments (Lada and Lada, 2003). In fact, massive stars are rarely observed to form in isolation (Zinnecker and Yorke, 2007). Moreover, a fraction of the massive stars that are observed to reside in isolation are likely to have been ejected from their parent cluster (see e.g. Schilbach and Röser 2008). Unfortunately, our understanding of massive star formation is limited, primarily due to the difficulty in observing these early stages of formation. The clustered environments of sites of on-going massive star formation contain large amounts of dust. As a result such objects suffer from high extinction, meaning that observational studies in the optical and IR are difficult. Moreover, these early evolutionary phases of massive star formation are both short lived and rare, considering the relatively short life time of massive stars in comparison to their significantly more abundant counterparts (low mass stars).

Typically, the onset of star formation is considered when a gas condensation (the core) begins to collapse within a larger subunit (a clump) of a dense molecular cloud (Williams, Blitz, and McKee, 2000). This leads to the formation of a protostar, which accumulates mass by accretion of the surrounding gas whilst at the same time losing mass through bipolar outflows (Zinnecker and Yorke, 2007). In low mass star formation, at the protostar stage, its core is not yet hot enough to start hydrogen (H) burning (at the onset of H burning, the object is generally considered to be on the main sequence). In massive star formation however, H burning can occur at the protostar stage (Kudritzki, 2002) and the process occurs within dense, compact clumps with typical $H_2$ column densities of $10^{23} - 10^{24}$ cm$^{-2}$, within giant molecular clouds. This embedded early phase of massive star formation can be sub-divided into different groups as discussed in Menten, Pillai, and Wyrowski (2005) and Tak and Menten (2005), namely IR dark clouds, hot molecular cores, hyper and ultra compact HII regions and compact classical HII regions.

Gas and dust surveys of various giant molecular clouds have revealed that such clumps are both very cold and very dense, with a typical temperature, T = 10 - 20 K, a molecular hydrogen density, $n_{H2} \sim 10^5$ cm$^{-3}$ over a region of $\sim$ 0.5 pc in diameter with a few $10^2 - 10^3$ $M_\odot$ of gas (Plume et al., 1997; Shirley et al., 2003; Garay et al., 2003; Motte et al., 2005). In the case of the IR dark clouds, these objects are seen in absorption against the mid-IR background emission and consist of a filamentary like structure containing numerous cold massive cores for the potential sites of massive star formation (Rathborne, Jackson, and Simon, 2006). It is thought that numerous massive star formation sites across a giant molecular cloud lead to the production of either bound OB stellar clusters, or unbound OB associations (Lada and Lada 2003; this discussion is returned to in Section 1.3.2).

The basic theory of massive star formation as outlined in the review of Zinnecker and Yorke (2007), follows a sequence of phases involving compression, collapse, accretion and disruption. Beginning with the formation of cold dense molecular cores (or filaments) within an existing giant molecular cloud, various regions within these cold dense cores collapse into optically thick protostellar embryos, which are supported by an outwards pressure force and have an initial mass of $\sim 10^{-3}$ $M_\odot$ (Larson, 1969; Bate, 2000). These embryos start to accrete their surrounding material, gaining mass as they evolve towards the main sequence (MS). Whilst low mass stars will not start burning hydrogen until they reach the MS, these high mass stars will begin H burning and also develop radiatively driven winds as they are accreting material and increasing in mass (Kudritzki, 2002). The stars stop gaining mass after a disruption of the parent cloud of gas and dust, which is brought about by the outflows of the surrounding massive stars and possible SNe within the same star-forming molecular cloud. For further information regarding the formation of massive stars, please refer to the annual review by Zinnecker and Yorke (2007).

### 1.1.3 The evolution of massive stars

Stellar evolution is dominated by changes to the nucleosynthetic make-up of stars. As these stars move off the main sequence (having exhausted all the hydrogen in their cores from nuclear fusion processes), they are pushed out of hydrostatic equilibrium, i.e. the inward

pull of gravity is no longer balanced by the outwards radiation and thermal pressure forces. Thus, their cores will contract and heat up, until heavier elements (e.g. He) are able to undergo nuclear fusion. In the process of exhausting a given element species, these objects become unstable and undergo changes to their luminosity, the size of their core and their outer envelope, i.e. altering their physical properties as they evolve towards a different evolutionary phase. It is the changes in their chemical composition (which in turn is dependant on their initial mass and their mass-loss rate), that are therefore, the driving force behind stellar evolution.

As mentioned previously, massive stars are those of mass larger than approximately eight solar masses. For low and intermediate mass stars (i.e. $0.6 \lesssim M \lesssim 8$ $M_{\odot}$), having exhausted their supply of H and He within their cores, they develop a degenerate carbon-oxygen core. Their core temperature does not reach that required in order to ignite nuclear carbon burning and these stars move onto the Asymptotic Giant Branch (AGB). At this evolutionary stage they are potentially subject to strong mass-loss episodes which shed their remaining envelope, leaving behind a carbon-oxygen (C-O) white dwarf. In comparison, massive stars will end their lives rather dramatically, possibly in the form of a supernova explosion, a long gamma-ray burst, a black hole or a neutron star. Above $\sim 8$ $M_{\odot}$, the cores of these massive stars can no longer be supported by electron degeneracy and as a result these cores evolve towards higher temperatures, fusing together heavier elements up until iron ($^{56}$Fe; at which point the fusion of heavier elements requires energy as opposed to producing it; Woosley, Heger, and Weaver 2002).

For stars of mass $M \gtrsim 8$ $M_{\odot}$, the amount of mass lost by their stellar winds throughout their entire lifetime (i.e. whilst also on the main-sequence) can have a profound effect on their subsequent evolution. Much work has been done to study the effect of mass-loss on the evolution of massive stars (see e.g. Chiosi and Maeder 1986; Maeder and Meynet 1987; Maeder and Meynet 2010). Figure 1.2 shows the evolutionary tracks in the HRD for stars of mass between 12 and 120 $M_{\odot}$, each of which have included the effect of stellar mass-loss and refer to objects with a metallicity $Z = 0.02$ (i.e. population I stars). This highlights the huge range in possible evolutionary scenarios for these massive stars, each of which are strongly dependant on their stellar mass. Thus, the mass-loss rate of these objects is a

hugely important parameter for the subsequent evolution of these objects.

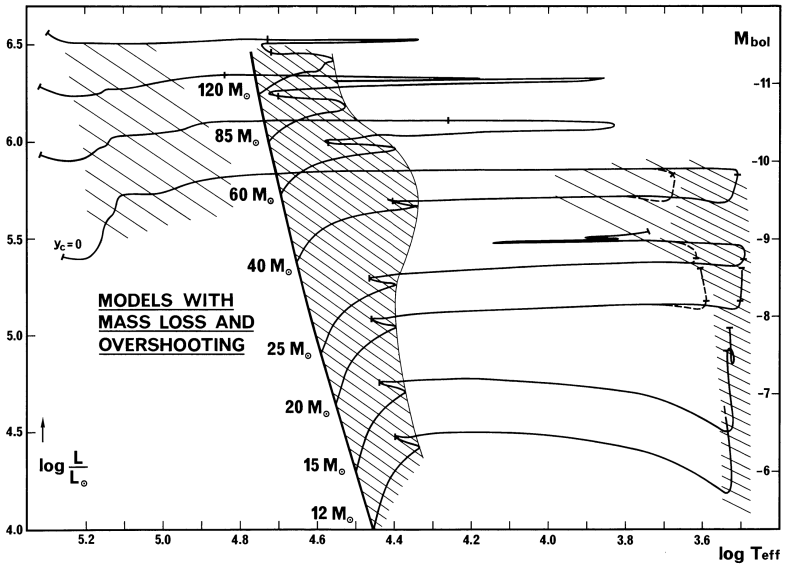



FIGURE 1.2: The evolutionary tracks of massive stars (12-120 $M_\odot$) as calculated by Maeder and Meynet (1987) including the effects of stellar mass-loss. The shaded regions are the 'long-lived' phases of the stars on the main-sequence and in the He burning RSG phase (shown to the right side of the HRD, for objects with $\log T_{eff} < 4.0$) and He burning WR phase (shown on the very left side of the HRD, for objects with $\log T_{eff} > 4.8$).

Possible evolutionary scenarios for these massive stars can be summarised as follows (where RSG: red supergiant; BSG: blue supergiant; WR: Wolf-Rayet star, SN: supernova and LBV: Luminous Blue Variable):

$M \lesssim 15\ M_\odot$ : MS (B-type) $\rightarrow$ RSG $\rightarrow$ SN II

$15 \lesssim M \lesssim 25\ M_\odot$ : MS (OB-type) $\rightarrow$ BSG $\rightarrow$ RSG $\rightarrow$ SN II[1]

$25 \lesssim M \lesssim 40\ M_\odot$ : MS (O-type) $\rightarrow$ BSG $\rightarrow$ RSG $\rightarrow$ WR $\rightarrow$ SN Ib[1]

$M \gtrsim 40\ M_\odot$ : MS (O-type) $\rightarrow$ BSG $\rightarrow$ LBV $\rightarrow$ WR $\rightarrow$ SN Ib/c[1]

A WR star is an evolved massive star that has been stripped of its hydrogen envelope during the RSG phase. These objects have strong stellar winds with typical mass-loss rates between $10^{-5} - 10^{-6}\ M_\odot yr^{-1}$, which have been found to depend strongly on their luminosity and chemical composition (see e.g. Nugis and Lamers 2000; Nugis and Lamers 2002). LBV stars are massive unstable objects found within the very upper region of the

[1]Supernova types: II = hydrogen rich; Ib = no hydrogen, some helium; Ib/c = no hydrogen, no silicon. All these types are core-collapse supernovae.

HRD. They are extremely luminous and close to the Eddington limit (i.e. the maximum luminosity a star can possess whilst maintaining hydrostatic equilibrium), have large mass-loss rates (up to $\sim 10^{-4}$ $M_{\odot}yr^{-1}$ ) and are observed to be variable both photometrically and spectroscopically (Humphreys and Davidson, 1994). The mass-loss rate tends to increase as a function of evolutionary phase for a given massive star and as a result their subsequent evolution becomes increasingly dependant on this parameter.

The evolutionary path of a given massive star is also affected by stellar rotation (see e.g. Heger, Langer, and Woosley 2000; Maeder and Meynet 2010), magnetic fields (Heger, Woosley, and Spruit, 2005), and if the object has a stellar companion (i.e. binarity; Podsiadlowski, Joss, and Hsu 1992; Sana et al. 2012). The latter is explored further in Chapter 5 of this thesis. The mass-loss rates of massive stars at a particular evolutionary phase (e.g. MS, RSG, BSG, LBV, WR) however, not only dictate their subsequent evolutionary phase, but also the type of core-collapse supernova associated with their eventual end point. All of these highlight the significance of such a parameter in the context of massive star evolution.

### 1.1.4 The role of massive stars

Massive stars are fundamental to the previous and on-going evolution of the Universe on both stellar, galactic and cosmological scales. As one of the primary sources of UV radiation, these objects are responsible for the re-ionisation of the Universe following the dark ages (Barkana and Loeb, 2001). Moreover, their ionising radiation leads to the production of countless HII regions readily observed throughout our own Galaxy and those further afield. Massive stars form the majority of heavy elements now present within our Universe through the fusion of successively heavier elements within their hot, dense cores. In addition, they play a crucial role in the chemical mixing and turbulent behaviour of the interstellar medium (ISM). Massive outflows of energy and material are constantly streamed from these objects via their stellar winds, whilst at the end of their lives violent gamma-ray bursts and supernova explosions deposit huge amounts of material into the surrounding

environment. These outflows and explosions interact with the surrounding ISM and ultimately provide the raw materials for the next generation of stars. The heavy elements found in dust grains, molecules and in both atomic and ionic form, act as a vital cooling mechanisms in the ISM and are ultimately a consequence of massive stars. Similarly, massive stars contribute to the cosmic ray population from the interaction of supernova shock fronts and Galactic magnetic fields, which in combination with the massive star UV radiation field, act as the principal sources of heating to the ISM (Zinnecker and Yorke, 2007). The processes involved in the formation of stars and planets are significantly dependant on the temperature and therefore any heating and cooling mechanisms. Massive stars must therefore be considered to significantly alter such star and planet forming processes (Bally, Moeckel, and Throop, 2005). On a larger scale, they impact the nucleosynthesis in galaxies, chemically enriching their surroundings whilst their various outflows can lead to the formation of Galactic 'super bubbles', thus affecting the physical and morphological structure of galaxies (see e.g. Kennicutt Jr 1998). Massive stars are either directly or indirectly related to numerous astrophysical subject areas and for these reasons, are often depicted as the 'cosmic engines' of our Universe.

## 1.2 Massive star winds

The vast majority of stars in our universe have a stellar wind. This wind can be simply described as an outflow of gas from the stellar surface. Most stars are similar in nature to that of our sun, which loses mass via its stellar wind at a rate of $\sim 10^{-14}$ $M_{\odot}yr^{-1}$ (Noerdlinger, 2008). These cool, late-type stars typically live for a long time (of the order of $10^{10}$ years) and have relatively little influence on their surrounding environment (at least whilst still on the main sequence, burning hydrogen in their cores). In contrast, massive, early-type stars are far less abundant within the universe, live far shorter lives ($\sim 10^{6}$ years for a 40 $M_{\odot}$ O-type star) and can lose mass at a rate of $\sim 10^{-6}$ $M_{\odot}yr^{-1}$. This huge amount of energy and material is blown radially outwards from the stellar surface and into the surrounding ISM.

Direct observational evidence of a stellar wind can be seen in the P-Cygni profiles of spectral lines. A P-Cygni line profile shows both absorption and emission components across the same spectral feature and is indicative of a gaseous envelope expanding radially outwards. The region of the wind where this expanding envelope lies between the star and the observer leads to the production of an absorption line that is blue-shifted relative to the star. This is combined with the other regions of the envelope which produce doppler broadened emission towards the red end of the spectrum relative to the rest wavelength of the spectral line. The blue absorption trough and red shifted emission lead to the characteristic P Cygni line profile as shown in Figure 1.3. UV resonance transitions that are connected to the ground state give rise to typical P-Cygni profiles (Kudritzki and Puls, 2000). Further spectral lines formed within a stellar wind include pure emission profiles or absorption lines refilled by wind emission due to recombination lines, a typical example of which in O-stars is the H$\alpha$ line.

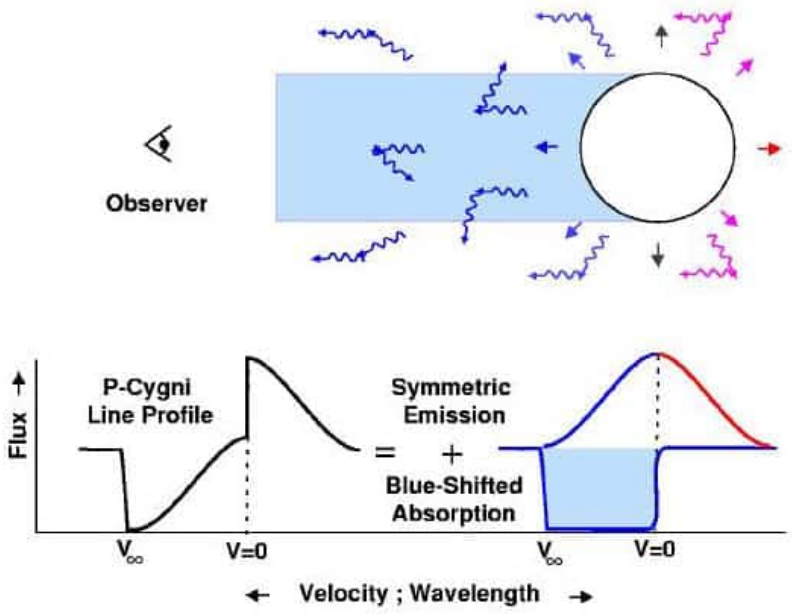



FIGURE 1.3: An example of a P-Cygni profile of a spectral line due to the radial outflow of gas via a stellar wind. Image from Daniel Kasen, http://supernova.lbl.gov/ dnkasen/.

The winds of massive stars are characterised and described by two fundamental parameters, the terminal velocity, $v_\infty$ and the mass-loss rate, $\dot{M}$. They are initiated at the stellar photosphere and are continuously accelerated radially outwards by a radiative driving

force due to the absorption of photons emitted from the photosphere (see Section 1.2.1). At an increasing distance from the stellar surface, the stellar radiation field will geometrically dilate, decreasing the radiative acceleration until it falls to zero, at which point the terminal velocity of the wind ($v_\infty$) will have been reached. For a homogeneous, stationary and spherically symmetric wind, the mass continuity equation yields the mass-loss rate as a function of the velocity field, $v(r)$ and the density distribution, $\rho(r)$, which are in turn a function of the radial distance $r$:

$$\dot{M} = 4\pi r^2 \rho(r) v(r) = constant. \tag{1.1}$$

This simply states that mass-loss rate through a spherical shell of radius $r$ that surround the star is constant. The velocity field in the supersonic region of the wind takes the form:

$$v(r) = v_\infty \left(1 - b\frac{R_\star}{r}\right)^\beta \tag{1.2}$$

where $R_\star$ is the radius of the stellar photosphere, $\beta$ describes the 'steepness' of the velocity law and the constant $b$ fixes the velocity at the onset of the wind (i.e. the inner boundary of the wind), which is normally taken to be the isothermal sound speed (Kudritzki and Puls, 2000). Whilst the terminal velocity of the wind can be directly measured, the mass-loss rate can be inferred from a number of techniques using observations at different wavelengths. Each of these have their own complexity and theoretical consideration and are briefly described in the following text.

1. **UV and Far UV Spectroscopy**

   P Cygni line profiles are readily observed in the UV spectrum of OB-type star winds and as touched on earlier are indicative of a gaseous outflow (i.e. a stellar wind). These profiles, along with UV absorption lines with violet wings (as seen in the UV spectrum of late O and early B-type main sequence stars) and UV emission lines (within Wolf-Rayet stars) can be compared to line synthesis models, for example those calculated using the Sobolev with Exact Integration (SEI) method (Lamers, Cerruti-Sola, and Perinotto, 1987; Haser et al., 1998). This diagnostic samples the

entire wind and is used to derive the column density of the particular atom in question, its distribution as a function of velocity, and the velocity law (the form of which is assumed). Determinations of the mass-loss rate, however, require further knowledge of the elemental abundances and the degree of ionisation of the ion responsible for each spectral line (Crowther et al., 2000).

2. **Optical and Near IR Spectroscopy**

   Alternatively, mass-loss information in early-type stars can been diagnosed from a treatment of the H$\alpha$ line ($\lambda$656.28 nm). Complex, nLTE radiative transfer codes such as CMFGEN (Hillier and Miller, 1998) or FASTWIND (Santolaya-Rey, Puls, and Herrero, 1997) are used to fit the optical and near IR stellar spectrum (including H$\alpha$, HeI, HeII lines) in order to derive the wind velocity law and mass-loss rate. Using the fit to the H$\alpha$ line probes the very inner wind region ($R \lesssim 3R_\star$). As a recombination line, it scales with the squared density ($\rho^2$) of the emitting material. Whilst observations of the H$\alpha$ line are straightforward, this technique leans heavily on the use of sophisticated nLTE model atmosphere codes since the line profile depends strongly upon the photospheric radiation field, the diffuse radiation field in the wind and the wind velocity law. An accurate knowledge of the H$\alpha$ line emitted from the photospheric radiation field is thus required, which for a given star has been found to be highly variable (Conti and Frost 1974; Ebbets 1980, Prinja. R, private communication).

3. **IR, mm and Radio Continua**

   Continuum emission from stellar winds across IR, mm and radio wavelengths occur due to free-free (Bremsstrahlung) 'thermal' radiation. Probing the intermediate ($10R_\star > R > 100R_\star$; IR, mm) and outermost ($R > 100R_\star$; radio) regions of the wind, the recombination line-formed continuum emission scales with the square of the wind density ($\rho^2$). Assuming spherical symmetry and a homogenous wind, mass-loss rates can be determined via the application of 'simple' analytical expressions such as those from Wright and Barlow (1975) and Panagia and Felli (1975). At these radial distances in the wind, the terminal velocity has been reached meaning no photospheric correction is required. However, care must be taken when considering the

ionisation state of the wind (see Section 4.2.4 for further information).

4. **X-Rays**

   Massive single stars have been readily observed to emit X-ray radiation. The effect is usually attributed to the presence of hydrodynamical shocks (a result of line de-shadowing instabilities, see Section 1.2.1) within their stellar winds, however the presence of a strong magnetic field has also been shown to promote X-ray emission in early type stars (Ud-Doula, Owocki, and Townsend, 2008). They have been shown to provide a good mass-loss rate diagnostic by analysing the effect of wind attenuation upon the shapes of their line profiles. Red-shifted photons emitted in the rear hemisphere of the wind are subject to a higher attenuation than those blue-shifted (emitted from the front hemisphere; Owocki and Cohen 2001). The column density of the wind (and therefore mass-loss rate) is directly proportional to the degree of this blue-shift and the asymmetry in the X-ray line profiles (Cohen et al., 2010; Cohen et al., 2011; Cohen et al., 2014). This diagnostic samples the entire wind region and in comparison to the UV line diagnostics, is relatively insensitive to the ionisation conditions of the wind.

### 1.2.1 Radiative line-driven wind theory

Observational evidence for mass-loss from hot stars became apparent during the late 1960s, when Morton (1967a) and Morton (1967b) discovered P Cygni like profiles of resonance lines in the far UV spectrum of three early supergiant stars. As the evidence mounted, it became clear that some mechanism must be responsible for these hot stars to lose mass via a stellar wind. The pioneering work of Lucy and Solomon (1970) and furthermore Castor, Abbott, and Klein (1975) (hereafter CAK), proposed a radiative line driving model whereby photospheric photons, each with a momentum $h\nu/c$, are scattered and absorbed by metal ions. The force imparted on the metal ions is shared with the more abundant hydrogen and helium species within the wind and must overcome the inwards force of gravity in order for them to accelerate the wind in a net outwards direction, resulting in significant mass-loss. The following aims to give a general overview of the theory behind radiative

line-driven winds, for a complete and more detailed description please refer to a review by Puls, Vink, and Najarro (2008b) and the references therein.

The stellar wind is accelerated by the transfer of momentum by line scattering. The net force is in the direction radially outwards from the star since the initial photons originate from a particular direction within the stellar photosphere. The subsequent re-emission after scattering/absorption with the metal ions is generally isotropic. The process is therefore symmetric with respect to the radial direction and the change in the direction angle of the photon, $\theta$ gives rise to the radial momentum transfer and hence the associated radiative line acceleration (Puls, Vink, and Najarro, 2008b). The total line acceleration in the radial direction $g_{rad}^{line}$ is obtained by integrating over all scatterings. This acceleration acts on the entire wind plasma if the momentum transferred to the metal ions can be sufficiently shared with the more abundant species within the wind. In a dense wind, coulomb coupling (via collisions) has been shown to successfully mediate this transfer of momentum to the more abundant H and He species (Puls, Vink, and Najarro, 2008b), allowing the entire wind to be treated as 'one' fluid.

To initiate a stellar wind, the total radiative acceleration must overcome the force of gravity at a particular point in the outer photosphere of the star. In order to include a full treatment of the outwards radiative force, $g_{rad}$, one must consider the contribution from the continuum, $g_{rad}^{cont}$ and that from optically thin and thick lines, $g_{rad}^{line}$. It is important to note here that the radiative force from all contributors is dependant upon the flux, $F$, radiated by the star and the opacity of the particles that scatter or absorb such radiation.

The continuum radiative force is due to the scattering of photons on free electrons, i.e. Thompson scattering, which can be written as:

$$g_{rad}^{cont} = \frac{\sigma_e F}{c} = \frac{\sigma_e L_\star}{4\pi r^2 c} \tag{1.3}$$

where $F$ is the flux radiated by the star, $\sigma_e$ is the Thompson scattering cross-section, $c$ is the speed of light, $L_\star$ is the total stellar luminosity and $r$ is the stellar radius. The line radiative force is the dominant contributor to the radiative acceleration (Puls, Vink, and

Najarro, 2008b). Depending on the strength of the transition, a line can be considered optically thick or optically thin. In the optically thick case, all the flux around a line of frequency $\nu_0$ is absorbed and depends upon the fraction of stellar flux within a bandwidth, $\Delta\nu$, centred around $\nu_0$ that is absorbed. This frequency bandwidth is derived under the Sobolev approximation (Sobolev, 1960), where it is assumed that there exists a large velocity gradient $dv/dr$, across a layer of width $\Delta r$, that does not change over a velocity interval, $V_t$, (i.e. the width of the line, approximately the mean thermal velocity of the photons):

$$\Delta r \approx \frac{V_t}{dv/dr} = L^{Sob} \tag{1.4}$$

This approximation holds in the supersonic region of the wind and also in the transonic regions assuming that the velocity gradient and opacity do not change over the line width (or velocity interval), $V_t$. The frequency bandwidth of flux absorbed by the line can therefore be written as:

$$\Delta\nu = \frac{V_t}{c}\nu_0 = \frac{\Delta r}{c}\left(\frac{dv}{dr}\right)\nu_0 \tag{1.5}$$

The total line acceleration in the optically thick case can therefore be written as:

$$g_{rad}^{thick} = \frac{F_{\nu_0}}{c^2}\nu_0\frac{dv}{dr} = \frac{L_\star\nu_0}{4\pi r^2 c^2}\frac{dv}{dr} \tag{1.6}$$

In the optically thin case, the line cannot absorb all of the flux emitted from the stellar photosphere at a frequency $\nu_0$. The amount of flux absorbed, is therefore proportional to the number of absorbing particles per unit volume such that the line acceleration becomes:

$$g_{rad}^{thin} = \frac{F_{\nu_o}}{c}\int_{line}\kappa_\nu\phi(\nu)d\nu \tag{1.7}$$

where $\kappa_\nu$ is the line absorption co-efficient and $\phi(\nu)$ is the profile function for absorption. Accounting for the contribution from both optically thin and optically thick lines, the total

radiative line force for a line at frequency $\nu_0$, can therefore be written as:

$$g_{rad}^{line} = \frac{F_{\nu_0}}{c^2}\nu_0 \left(\frac{dv}{dr}\right)\left(1 - e^{-\tau_{Sob}}\right) \tag{1.8}$$

where the profile for the absorption profile $\phi(\nu)$ has been approximated to a delta function and $\tau_{Sob}$ is the optical depth in the Sobolev approximation, i.e.

$$\tau_{Sob} = \kappa_\nu \frac{c}{\nu_0}\left(\frac{dr}{dv}\right) \tag{1.9}$$

Equation 1.8 gives the radiative line acceleration under the Sobolev approximation for a line at frequency $\nu_0$. In the optically thick case, when $\tau_{Sob} >> 1$, Equation 1.8 reduces to Equation 1.6. Similarly in the case of optically thin lines ($\tau_{Sob} << 1$), Equation 1.7 is reproduced. In order to derive the total radiative line acceleration, Equation 1.8 must be summed over all contributing lines. Following the work of CAK, this can be expressed in terms of a force multiplier, $M(t)$ and the radiative acceleration due to electron scattering, $g_{rad}^{Th}$:

$$g_{rad}^{line}(r) = g_{rad}^{Th} M(t) = \frac{\sigma_e^{ref} L_\star}{4\pi r^2 c} M(t) \tag{1.10}$$

where $\sigma_e^{ref}$ is a reference value of the Thompson scattering cross-section (CAK used a value of 0.325 cm$^2$ g$^{-1}$). The force multiplier, $M(t)$ is a function of the parameters $k$, the number of spectral lines and $\alpha$, which describes the distribution of strong to weak lines (see Puls, Springmann, and Lennon 2000 for further details):

$$M(t) = kt^{-\alpha} \tag{1.11}$$

where $t$ is the optical depth of a line with an opacity equal to the Thompson scattering (Puls, Vink, and Najarro, 2008b):

$$t = \sigma_e V_t \rho \left( \frac{dr}{dv} \right) \tag{1.12}$$

Having derived the line and continuum radiative line force, one can now construct the equation of motion to describe a hot star wind. Under the assumption of a stationary, homogeneous and spherically symmetric wind this reads:

$$v \frac{dv}{dr} = -\frac{GM_\star}{r^2} - \frac{1}{\rho} \frac{dP}{dr} + g_{rad}^{tot} \tag{1.13}$$

This equation is further simplified by neglecting the gas pressure term on the right hand side (since the enthalpy term in the energy equation is significantly smaller than the potential and kinetic terms; Puls, Vink, and Najarro 2008b) and including the contribution from the radiative continuum force (Equation 1.3), i.e. $g_{rad}^{tot} = g_{rad}^{cont} + g_{rad}^{line}$:

$$v \frac{dv}{dr} = -\frac{GM_\star}{r^2} + \frac{\sigma_e L_\star}{4\pi r^2} + g_{rad}^{line} \tag{1.14}$$

The second term on the right hand side describing the contribution from the continuum radiation can be re-written in terms of the Eddington factor, $\Gamma_e GM_\star = L_\star \sigma_e / 4\pi c$. Furthermore, using the expression for the effective mass $M_{eff} = M_\star(1 - \Gamma_e)$, Equation 1.14 simplifies to:

$$v \frac{dv}{dr} = -\frac{GM_{eff}}{r^2} + g_{rad}^{line} \tag{1.15}$$

Inserting the expression for the total line radiative acceleration, multiplying each term by $r^2$ and substituting in the mass continuity equation in the final term on the right hand

side obtains:

$$r^2 v \frac{dv}{dr} = -GM_{eff} + \frac{\sigma_e L_\star}{4\pi c} k \left( \frac{\sigma_e V_t \dot{M}}{4\pi r^2 v} \frac{dr}{dv} \right)^{-\alpha} \tag{1.16}$$

By then defining the parameter $y = r^2 v dv/dr$, absorbing the remaining constants on the final term into the parameter $f$, and re-arranging, the equation of motion becomes:

$$f y^\alpha - y - GM_{eff} = 0 \tag{1.17}$$

By differentiating and solving for $f$, one can then obtain $y$ by a direct substitution back into Equation 1.17:

$$f = \frac{1}{\alpha} y^{(1-\alpha)} \qquad \text{and} \qquad y = GM_{eff} \left( \frac{\alpha}{1-\alpha} \right)$$

By inserting the solution to $y$ back into its original definition and integrating the left and right hand sides between 0 and infinity and $R_\star$ and infinity respectively, one obtains the equation of motion in its integral form:

$$\int_0^\infty v dv = GM_{eff} \left( \frac{\alpha}{1-\alpha} \right) \int_{R_\star}^\infty r^{-2} dr \tag{1.18}$$

By defining the escape velocity as $v_{esc} = (2GM/R_\star)^{0.5}$ and integrating Equation 1.18, one obtains the relationship between the escape and terminal velocity as:

$$\left( \frac{v_\infty}{v_{esc}} \right) = \left( \frac{\alpha}{1-\alpha} \right)^{0.5} \tag{1.19}$$

If considering the velocity law as previously stated in Equation 1.2, this result from CAK predicts a value of $\beta = 0.5$. Moreover, an expression for the mass-loss rate of the star drops

out of the solution to $f$ as:

$$\dot{M} = \left(\frac{\sigma_e}{4\pi c}\right)^{1/\alpha} \frac{4\pi}{\sigma_e V_t}(\alpha k L_\star)^{1/\alpha} \left(\frac{1-\alpha}{GM_{eff}}\right)^{\frac{\alpha-1}{\alpha}} \tag{1.20}$$

This is a simplified solution from Kudritzki et al. (1989) that is valid in the limit of a small sound speed ($a << v_{esc}$), applies only under the assumption that the stellar radiation is radially streaming from a point source (Owocki, 2014), and assumes that the parameters $k$ and $\alpha$ from the force multiplier equation (1.11) are constant. For a more elaborate study that undergoes the relaxation of such parameters, please refer to Kudritzki (2002).

The above model to radiative line driven winds is subject to a number of assumptions which besides those aforementioned are listed here:

1. A core-halo structure has been adopted, meaning continuum formation in the wind has been neglected (important in the case of dense winds; Puls, Vink, and Najarro 2008b).

2. Multi-line effects have been neglected, meaning each line only undergoes one interaction with the unattenuated stellar radiation field, leading to an overestimation of the radiation force (Friend and Castor, 1983; Abbott and Lucy, 1985; Gayley, 1995; Puls, Vink, and Najarro, 2008b).

3. Effects such as wind-blanketing and back-warming due to the effect the wind has on the emergent radiation field have not been taken into account.

Since the work of CAK throughout the 1970s, major improvements have been made in computing the wind structure of hot stars (see Friend and Abbott 1986; Pauldrach, Puls, and Kudritzki 1986; Pauldrach 1987; Pauldrach et al. 1994; Pauldrach, Hoffmann, and Lennon 2001; Kudritzki 2002; Krtička and Kubat 2006). These include a departure from Local Thermodynamic Equilbrium (LTE) and the inclusion of millions of lines from over 150 ions. The improvements have led to reasonably good agreements to observed values (Abbott, 1978; Abbott, 1982; Cassinelli, 1979), however the majority of which still depend crucially on the use of the Sobolev approximation (Owocki, 2014).

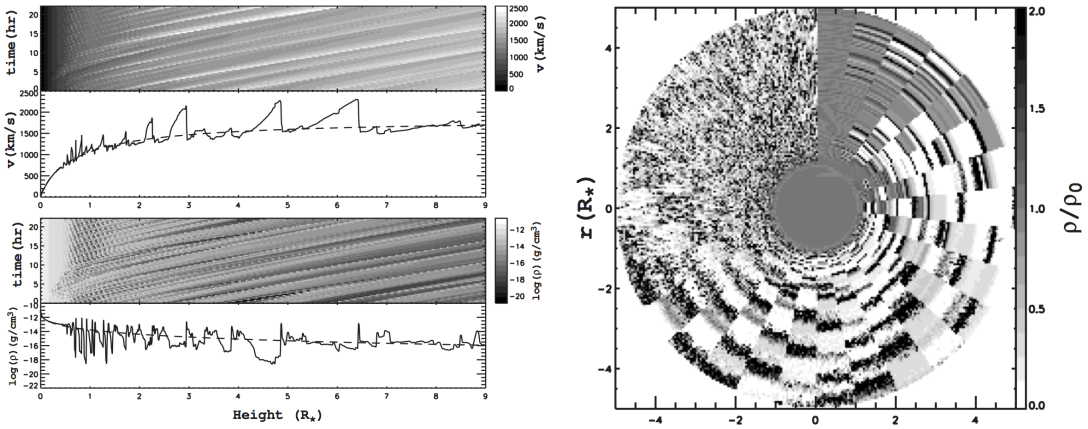



FIGURE 1.4: Plots to highlight the results from 1-D (left) and 2-D (right) SSF models (as discussed in Dessart and Owocki 2005; Owocki 2014) of the line-deshadowing instability in hot stars. The line plots on the left panel show the velocity (upper) and density (lower) variations as a function of stellar radii at a fixed time. The grey-scale plots above them show their respective evolution as a function of time. For a 2-D hydrodynamical simulation, the right plot represents the density variations throughout the stellar wind shown in chunks of time, stacked clockwise every 12° in intervals of 4000s. Both plots have been taken directly from Owocki (2014).

Stellar wind models that relax the Sobolev approximation show that the stellar wind is subject to a strong 'line de-shadowing instability' (LDI) for those perturbations in the wind velocity that are less than or close to the Sobolev length. This effect is thought to make the winds highly structured and variable (MacGregor, Hartmann, and Raymond, 1979; Owocki and Rybicki, 1984; Owocki and Rybicki, 1985; Owocki, 2014). In the wind, line de-saturation occurs due to the Doppler shift imparted on the lines due to the wind acceleration. This not only affects the line-driving mechanism, but leads to strong LDI. Small-scale velocity perturbations at the onset of the wind are thought to grow due to the de-shadowing of the line caused by the extra Doppler shift gained from the initial velocity perturbation (Owocki and Rybicki, 1984). Various one and two dimensional Smooth Source Function (SSF) simulations have predicted extensive velocity and density structures within hot star winds, such as those pictured in Figure 1.4 (Owocki and Rybicki, 1984; Owocki and Rybicki, 1985; Dessart and Owocki, 2005). The prediction of such structured winds has a profound effect upon various wind diagnostics and have led to an uncertainty within current mass-loss estimates. The following Section (1.2.2) will explore the various density and velocity structures thought to occupy the winds of massive stars and provide an indication

of the effect they can have on the observational stellar wind diagnostics.

### 1.2.2 Structure in the winds of hot stars

For almost half a century, it has been suggested that structure (or density inhomogeneities) exists within the winds of early-type stars. It began with the work of Osterbrock and Flather (1959), who postulated that clumps or extreme density fluctuations could explain the discrepancy between the observed and predicted fluxes for points within an HII region of the Orion nebulae. This soon moved into the realms of hot star winds where Lucy and White (1980) postulated that 'small-scale' clumps within the wind of $\zeta$ Puppis could account for their observed X-ray luminosity. In a similar fashion, Abbott, Beiging, and Churchwell (1981) investigated the effect such (small-scale) wind structure would have on the thermal free-free emission from massive star winds. Additionally, a different type of wind structure (large-scale) was used to explain the variability across UV P Cygni line profiles of massive stars (Howarth and Prinja, 1989).

As structured wind models emerged in order to explain different observations across the electro-magnetic spectrum, it soon became apparent that two distinct types of wind structures existed: small-scale structure, also referred to as wind 'clumping', and large-scale structure known as co-rotating interaction regions (CIRs). Whilst theoretical work by MacGregor, Hartmann, and Raymond (1979), Owocki and Rybicki (1984), Owocki, Castor, and Rybicki (1988), Feldmeier, Puls, and Pauldrach (1997), and Runacres and Owocki (2002) began to incorporate wind-clumping into their simulations of hot star winds, the first clear observational evidence of clumping was provided by Eversberg, Moffat, et al. (1998) who found moving structures within the HeII$\lambda$4686 wind line profile of $\zeta$ Pup. Clumping has been further used to explain the observed discrepancies between $\rho$ (UV resonance lines) and $\rho^2$ (H$\alpha$/radio) mass-loss diagnostics (see above for further details). Due to the strong dependance of $\dot{M}$ diagnostics upon the degree and type of wind structure present within early type stars, both small- and large-scale wind structures are explored in turn below.

#### 1.2.2.1 Small-scale structure - clumping

Wind clumping describes small-scale density in-homogeneities that re-distribute matter into enhanced density clumps of a particular size, $l(r)$. However, clumping can be further characterised depending upon the optical depth or size of the 'over-dense' clump itself. Micro-clumping describes the optically thin case, whereby the size of the clump is smaller than the photon mean free path of the matter-light interaction (Puls, Vink, and Najarro, 2008b). Macro-clumping on the other hand refers to those clumps that are optically thick at certain frequencies (and therefore to certain spectral lines), where their size is larger than that of the photon mean free path (see Oskinova, Hamann, and Feldmeier 2007).

In the micro-clumping approach, a leading assumption states that the inter-clump material is void, i.e. a vacuum exists in the regions between the over-dense clumps of wind material. The average density of the wind, $\langle\rho\rangle = \dot{M}/4\pi r^2$ is expressed in terms of the volume filling factor, $f_{vol}$ and the density of the clump material, $\rho_c$ as:

$$\langle\rho\rangle = f_{vol}\rho_c \quad , \quad \langle\rho^2\rangle = f_{vol}(\rho_c)^2$$

Moreover, Owocki, Castor, and Rybicki (1988), who developed time dependant models of radiatively driven winds, introduced the clumping factor, $f_{cl}$ to measure the over-density of the clumps, that is simply the inverse of $f_{vol}$ and defined as:

$$f_{cl} = \frac{\langle\rho^2\rangle}{\langle\rho\rangle^2} \quad , \quad f_{cl} = f_{vol}^{-1} \quad , \quad \rho_c = f_{cl}\langle\rho\rangle$$

The corresponding optical depth, $\tau$ is expressed in terms of the opacity and the volume filling factor:

$$\tau = \int \kappa f_{vol} dr = \int \kappa_C(f_{cl}\langle\rho\rangle) f_{vol} dr \tag{1.21}$$

calculated with a reduced path length ($f_{vol}dr$) and where $\kappa_C(f_{cl}\langle\rho\rangle)$ is the opacity as evaluated inside the clump. As a result, the opacity for linearly dependant density processes

(e.g. UV resonance line profiles) is the same as in the smooth wind, homogeneous case. For density squared dependant processes, however, the opacity (and therefore emissivity) scales with $f_{cl}$, resulting in a dependance upon the mass-loss rate of $\dot{M} \sim \sqrt{f_{cl}}$.

Originally treated as being constant throughout the entire wind, $f_{cl}$ was first thought to vary as a function of radial distance from the star by Nugis, Crowther, and Willis (1998). The authors attempted to reconcile mass-loss rates inferred from IR, mm and radio measurements of Wolf-Rayet (WR) stars, by suggesting a radial stratification of the clumping factor. This was further investigated by Puls et al. (2006), who simultaneously modelled H$\alpha$, IR, mm and radio emission from 19 different O-type stars with well defined parameters. For dense winds, it was found that considerable amounts of clumping were found close to the stellar surface ($R < 3R_\star$), which would remain until the outer wind regions ($R \sim 100R_\star$), where it would decrease. For stars with weaker winds however, Puls et al. (2006) found a better consistency between $f_{cl}^{inner}$ and $f_{cl}^{outer}$. Taking the constraints on the radial stratification of the clumping factor found in the literature, Najarro et al. (2008) suggested the following clumping law:

$$f_{vol}(r) = CL1 + (1 - CL1)e^{\frac{-v(r)}{CL_2}} + (CL4 - CL1)e^{\frac{-(v_\infty - v(r))}{CL3}} \tag{1.22}$$

where CL1 and CL4 are the maximum and outer wind region values of the volume filling factor respectively, and CL2 and CL3 are velocity terms that define the location in the wind at the onset of clumping and when the clumping changes to the value of CL4, respectively. For reference, this clumping law is plotted for a range of values of CL2 and CL3 in Figure 1.5.

The mass-loss rates inferred from $\rho^2$ diagnostics (H$\alpha$; IR, mm, radio continua) scale with the square root of the clumping factor. As a result, the inclusion of micro-clumping at a given $f_{cl}$ will seek to lower the mass-loss rate for a given diagnostic (e.g. radio flux). Micro-clumping has been used to reconcile the disparity found by Fullerton, Massa, and Prinja (2006) between mass-loss rates inferred from $\rho^2$ diagnostics and those from the PV resonance lines ($\propto \rho$) as highlighted in Figure 1.6. However, when taking into account macro-clumping (the effect of which is also referred to as porosity), Oskinova, Hamann, and

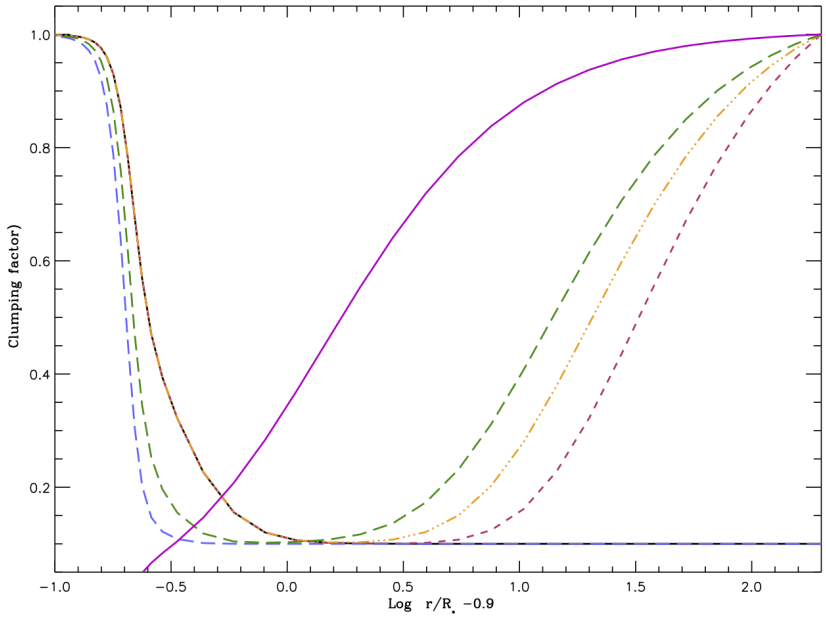



FIGURE 1.5: A model of the radial stratification of the clumping factor, $f_{vol}(r)$ for different values of CL2 and CL4 (represented by the dashed lines) as given in Equation 1.22. The solid pink line represents the velocity structure in units of $v_\infty$. The image (and indeed the clumping law) has been taken from Najarro et al. (2008).

Feldmeier (2007) and Šurlan et al. (2013) have been able to reconcile the disparity between H$\alpha$ and PV mass-loss diagnostics for O-type stars (as shown in Figure 1.6), without the need for a high clumping factor.

In the macro-clumping approach, it is assumed that clumps of size $l(r)$ and mass $m_C(r)$ are larger than the photon mean free path of the matter-light interaction and are hence optically thick at some frequency $\nu$. As such, the process depends not only on the size of the clumps, but their distribution and geometry. Defining $L$ as the average separation between the clumps of average size $l$, the volume filling factor can be defined as:

$$f_{vol} = \left(\frac{l}{L}\right)^3 = \frac{1}{f_{cl}} \tag{1.23}$$

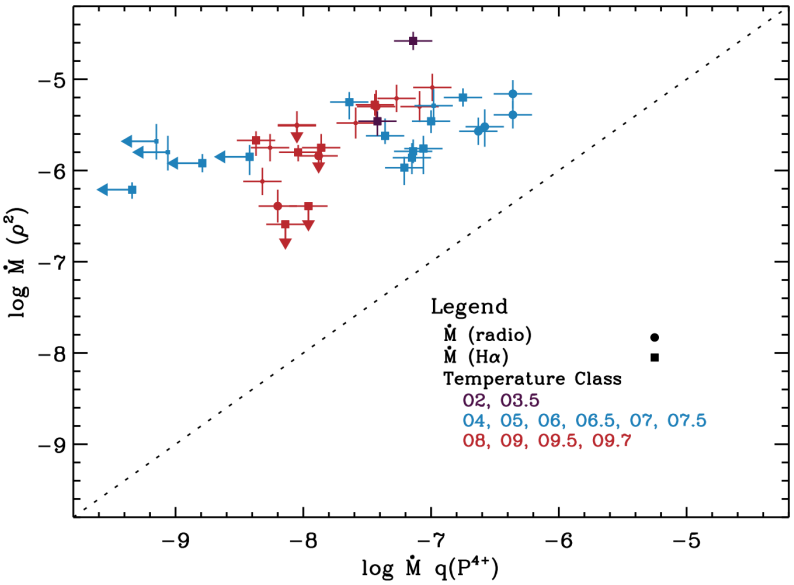



FIGURE 1.6: A comparison of the derived $\dot{M}$ values from $\rho^2$ diagnostics with those derived using the PV resonance line for a sample of 40 Galactic O-type stars. Image taken from Fullerton, Massa, and Prinja (2006).

Assuming each clump has a density, $\rho_C \approx m_C/l^3$, the optical depth across a clump with opacity $\kappa_C$ can be written as:

$$\tau_C \approx \kappa_C \rho_C l = \kappa_C \frac{m_C}{l^2} \tag{1.24}$$

If the clumps are optically thick, then their effective cross section is simply $\sigma_{eff} = l^2$. However, if expressed more generally, $\sigma_{eff}$ should scale with the fraction of radiation attenuated by the clump, i.e. $\sigma_C = l^2(1 - e^{-\tau_C})$. The effective opacity of the clumps can therefore be written as:

$$\kappa_{eff} = \frac{\sigma_C}{m_C} \approx \frac{l^2}{m_C}(1 - e^{-\tau_C}) = \frac{\kappa_C}{\tau_C}(1 - e^{-\tau_C}) \tag{1.25}$$

Note that in the optically thin case ($\tau_C \rightarrow 1$), the $\kappa_{eff} = \kappa_C$ as in the micro-clumping approximation. In the optically thick case however ($\tau_C >> 1$), the effective opacity is reduced by a factor of $1/\tau_C$. Owocki, Gayley, and Shaviv (2004) introduced the porosity length, $h$ which quantifies the photon mean free path in a medium of optically thick clumps.

This porosity length is defined as $h = L^3/l^2$ and can be related to the optical depth in Equation 1.24 and hence the effective opacity can be re-expressed as:

$$\kappa_{eff} = h^{-1}(1 - e^{-\tau_C}) \tag{1.26}$$

The net effect of macro-clumping on the optically thick lines due to optically thick clumps is that their opacity is significantly reduced. This has the effect of weakening the resonance line profile and hence explaining the lower mass-loss rates derived from optically thick UV resonance lines in O-type stars (as found in Fullerton, Massa, and Prinja 2006). Oskinova, Hamann, and Feldmeier (2007) modelled the spectrum of the O4 I $\zeta$ Pup including macro-clumping (porosity) effects to find that the optically thin H$\alpha$ line is not affected by wind porosity. They did, however, find that the optically thick PV resonance doublet became significantly weaker, concluding that empirical mass-loss rates from recombination line diagnostics were in no need of a downwards revision.

It is important to note the impact that wind structure has on the specific mass-loss diagnostic being used. In order to obtain a reliable $\dot{M}$, one must first understand the effect that any possible wind structure has on a given diagnostic, and secondly the extent to which the object is affected by such structure (i.e. the value of $f_{cl}$). Moreover, to fully understand and indeed correctly model the stellar winds of massive stars, a multi-wavelength approach is of absolute necessity and allows a derivation of the mass-loss rate and constraints to be made on the radial stratification of the clumping at multiple points throughout the entire wind region.

#### 1.2.2.2 Co-rotating interaction regions - CIRs

Additional to small-scale wind structures such as micro and macro clumping, a considerable amount of observational evidence in the form of Discrete Absorption Components (DACs) suggest the presence of large-scale structures within the winds of hot stars. DACs can be seen within the absorption troughs of unsaturated UV P-Cygni line profiles and are due to variations in the optical depth along the line of sight from within the stellar wind

(Howarth and Prinja, 1989). An example of a dynamic spectrum of the O4If(n) stars ζ Pup, displaying these DAC features can be seen in Figure 1.7. DACs have been observed to behave systematically, accelerating towards the blue wings of the line profile over the timescales of a few days, becoming narrower as they approach the asymptotic velocity. They have been observed in O- and early B-type stars (Howarth and Prinja, 1989), late B-supergiants (Bates and Gilheany, 1990), a WN7 star (Prinja and Smith, 1992) and even in the central stars of planetary nebulae (PN; Prinja, Massa, and Cantiello, 2012).

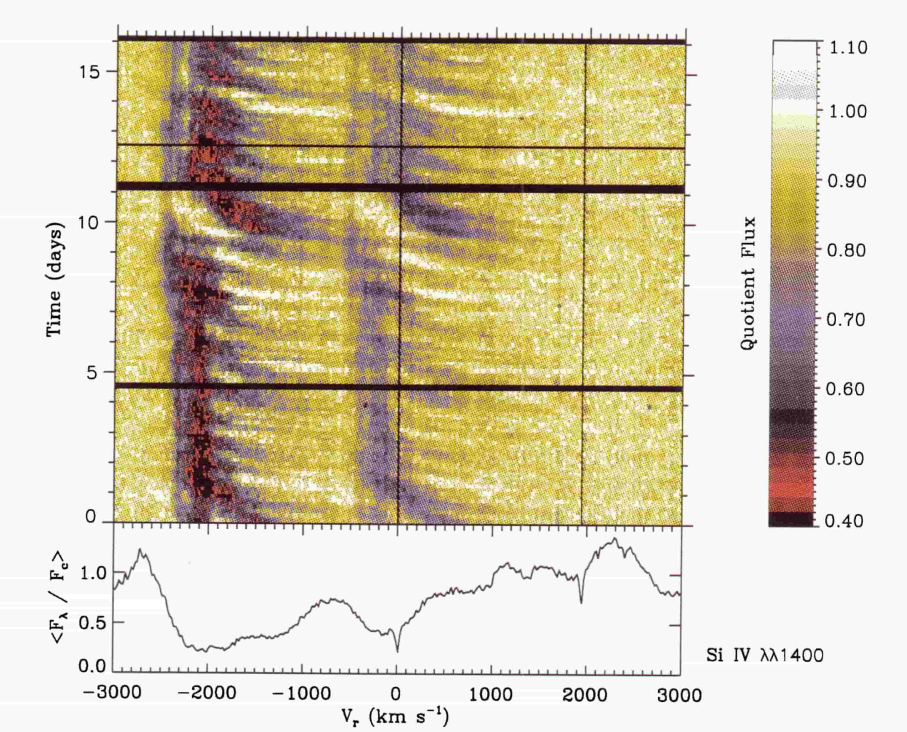



FIGURE 1.7: A dynamic spectrum of the Si V λ1393, 1402 resonance doublet for the O4If(n) star ζ Pup as observed as part of the IAU MEGA campaign (Massa et al., 1995). The plot shows how discrete absorption components (DACS) migrate along the line profile, accelerating towards the blue wings, indicative of large-scale wind structure (CIRs).

The rate at which the DACs reappear and their acceleration across the absorption profile has been correlated with the rotational period of the star, i.e. $v \sin i$ (Prinja, 1988; Henrichs, Kaper, and Zwarthoed, 1988; Kaper et al., 1999). Moreover, this acceleration has been found to be less than that of the wind material, as diagnosed from the shapes of saturated wind lines (Prinja, 1994). To further explain why different features have been observed to propagate at different accelerations, Mullan (1984) proposed that a co-rotating interaction region (CIR), similar to the structures seen in the solar wind, can give rise to

such absorption features. It was not until the 2-D hydrodynamical simulations of Cranmer and Owocki (1996) that the CIR scenario was properly investigated. It was proposed that photospheric disturbances in the azimuthal direction, caused by a local increase in the line force due to a bright equatorial spot, could induce such a wind structure. The exact cause of the bright stellar spots is still uncertain, although mechanisms such as non-radial pulsations and localised magnetic fields have been proposed (Owocki, Cranmer, and Fullerton, 1995; Fullerton et al., 1997; Puls, Vink, and Najarro, 2008b).

The addition of locally applied line acceleration generates streams of material of higher density and of lower velocity in comparison to the mean wind flow. These generated streams propagate radially outwards as the star rotates as shown in Figure 1.8. Where the supersonic mean wind flow collides with the inside edge of these streams, a co-rotating interaction region is formed. A non-linear signal is then sent back towards the stellar surface after the collision, forming a sharp discontinuity that propagates in the radial velocity gradient known as an 'Abbott-kink' (see Feldmeier and Shlosman 2000). This 'kink' moves slowly radially outwards with respect to the stellar frame of reference and it is these kinks in conjunction with the shallow velocity gradients that Owocki, Cranmer, and Gayley (1996) showed to be the origin of the observed DACs.

The effect of the CIRs upon the mass-loss rates of massive stars has been shown to be small (Lobel and Blomme, 2008). They do, however, have a significant impact on the wind lines by altering their optical depths. Whilst optically thick wind structures have been shown to alter the mass-loss diagnoses by UV spectral lines (Prinja and Massa, 2010), and especially so in the case of small-scale structures (i.e. macro-clumping Oskinova, Hamann, and Feldmeier 2007; Šurlan et al. 2013), it is not clear to what extent large-scale optically thick structures affect the resulting wind profiles. At any given time, CIRs can affect a large velocity range, however they have yet to be modelled self-consistently alongside small-scale wind structure, meaning their effect on the observed wind profiles is not fully understood.

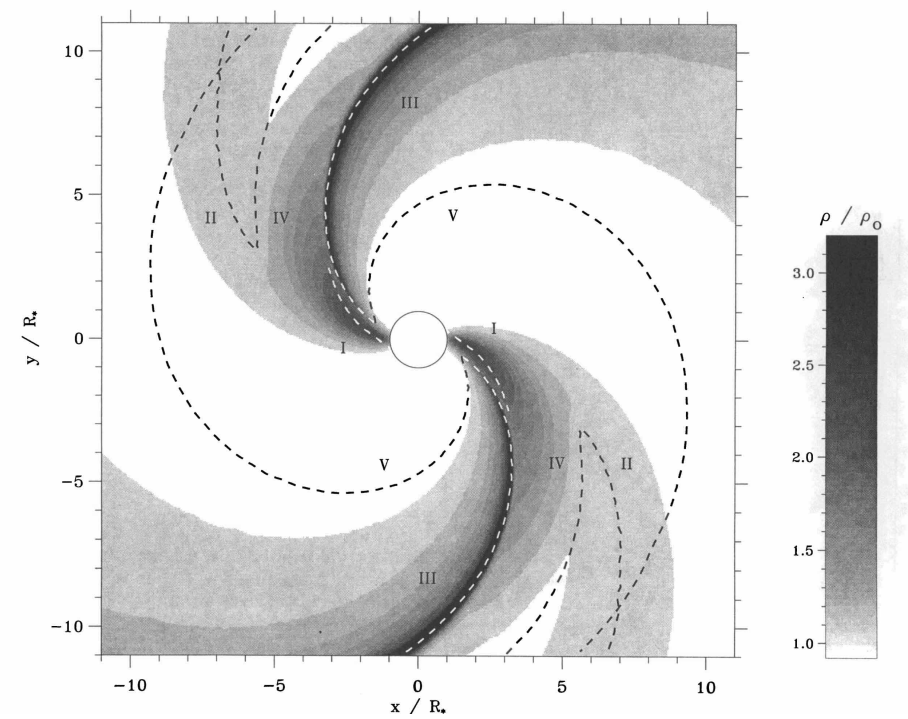



FIGURE 1.8: A 2-D greyscale plot of a CIR model. The bright equatorial spot of highlighted at the position of I, the dashed line labelled as III shows the compressed CIR, and V highlights the post shock 'Abbott-kink'. The model and image were taken from the work of Cranmer and Owocki (1996).

## 1.3 Massive stellar clusters

This section widens the discussion from massive stars to massive stellar clusters (MSC). Hosting numerous stellar objects throughout all phases of evolution, Young Massive Clusters (YMCs) are the key sites of massive star formation and therefore the best 'playgrounds' in which to study the formation and evolution of massive stellar populations. As an ideal target for star formation studies across the Hertzspring-Russel diagram (HRD), they show a direct link between the chemistry and dynamics of stars and their local environment. The massive star content within YMCs are thought to influence galaxy formation and evolution throughout cosmological times, whilst also having a direct impact upon the re-ionisation of the inter-galactic medium.

### 1.3.1 Fundamental properties of massive stellar clusters

Qualitatively, massive stellar clusters can be described as a large population of stars with varying properties, most notably of different masses. In order for a stellar cluster to be defined as 'massive', it must have a mass larger than $\sim 10^3$ $M_\odot$. The majority of stars form in clustered environments (e.g. see Lada and Lada, 2003; Gutermuth et al., 2009). In particular, massive (OB) stars are often found within stellar clusters, as seen by the clustering of O-type stars within the Milky Way by Parker and Goodwin (2007). Furthermore, Wit et al. (2005) found that just 4 ± 2 % of O-stars can be considered to have formed outside a clustered environment, whilst further studies (Gvaramadze and Bomans, 2008; Schilbach and Röser, 2008; Pflamm-Altenburg and Kroupa, 2010) suggest that a fraction of that 4% are runaway stars. In young-starburst galaxies[2], it has been found that at least 20% of the UV light (but more likely all), has been found to come from YMCs (Meurer et al., 1995). Merging galaxies contain huge numbers of YMCs and Fall, Chandar, and Whitmore (2005) found a median cluster age within the Antennae galaxy of just 100 Myr, concluding that in this case at least 20% and possibly all its stars form in clustered environments.

Key measurable quantities of a MSC include its mass, age and luminosity, yet many parameters have been used to characterise stellar clusters. There are generally considered to be three distinct groups of stellar cluster. Open clusters (OCs), which represent systems that are gravitationally unbound that will eventually dissolve into the ISM. YMC, which are young, gravitationally bound systems with sites of on-going star formation and finally globular clusters (GCs) that describe older, gravitationally bound systems with no sites of active star formation. Generalised quantities that give approximate distinctions between YMC, OC and GCs can be found in Table 1.1, taken from Portegies Zwart, McMillan, and Gieles (2010).

The virial radius, $r_{vir}$ is a theoretical prediction of a characteristic length scale that assumes the system is in virial equilibrium, i.e. the potential energy of the system is equal to twice the total kinetic energy of the system. The virial radius links the total mass of the system

[2] A starburst galaxy can be defined a galaxy that is converting its gas into stars at an exceptionally high rate, typically in comparison to the average star formation rate observed in most other galaxies

TABLE 1.1: A comparison between the fundamental parameters of the three outlined stellar clusters as given by Portegies Zwart, McMillan, and Gieles (2010).

| Cluster | Age (Gyr) | M ($M_\odot$) | $r_{vir}$ (pc) | $\rho_c$ ($M_\odot$ $pc^{-3}$) | Z ($Z_\odot$) | Location | $t_{dyn}$ (Myr) | $t_{rh}$ (Myr) |
|---|---|---|---|---|---|---|---|---|
| OC | $\lesssim 0.3$ | $\lesssim 10^3$ | 1 | $\lesssim 10^3$ | $\sim 1$ | Disk | $\sim 1$ | $\lesssim 100$ |
| GC | $\lesssim 10$ | $\gtrsim 10^5$ | 10 | $\gtrsim 10^3$ | $< 1$ | Halo | $\gtrsim 1$ | $\gtrsim 1000$ |
| YMC | $\lesssim 0.1$ | $\gtrsim 10^4$ | 1 | $\gtrsim 10^3$ | $\gtrsim 1$ | Galaxy | $\lesssim 1$ | $\lesssim 100$ |

to its total potential energy, a quantity that cannot be directly observed and must be derived from the stellar masses and positions within a N-body simulation. Open and young massive clusters are often found to have a similar values of $r_{vir}$, whilst older GC systems have a much larger virial radius by approximately an order of magnitude (see Table 1.1). The dynamical time $t_{dyn}$, and the relaxation time $t_{rh}$ are important quantities when describing any self-gravitating system. The time taken for a typical star to cross the system is described by $t_{dyn}$ which, if the system is in virial equilibrium, equates to be

$$t_{dyn} = \left(\frac{GM}{1.58 r_{vir}^3}\right)^{-1/2} \approx 2 \times 10^4 yr \left(\frac{M}{10^6 M_\odot}\right)^{-1/2} \left(\frac{r_{vir}}{1pc}\right)^{3/2}, \tag{1.27}$$

as was shown by Spitzer Jr and Hart (1971). This is generally found to be larger for older, massive clusters such as GCs as opposed to younger star forming clusters (see Table 1.1). The relaxation time, $t_{rh}$ can be thought of as the time taken for a typical cluster member to lose all sense of its original trajectory through various stellar encounters and is often much greater than the cluster's dynamical timescale. $t_{rh}$ can vary significantly between the core and outer regions of the system due to the radial density profile of the cluster. For the remaining parameters given in Table 1.1, $\rho_c$ describes the density of the cluster core and $Z$ describes the metal content of the cluster.

An unbound or open cluster can be more specifically defined as a stellar system whose total energy, (i.e. gravitational and kinetic energy) is positive, such that the internal motions of its members are too large to be gravitationally confined by the baryonic material found within the cluster boundary (Lada and Lada, 2003). Similarly, the distinction can be made for bound clusters whereby their total energy is negative. Open clusters form a large

population of all varieties of stellar systems with Lada and Lada (2003) suggesting that only $\sim$ 10% of stars are found in gravitationally bound systems within 10 Myr of their birth. Midway through the 20th century, another subgroup of clusters emerged known as stellar associations, a term originally founded by Ambartsumian (1954) and that was later described as a low density group of stars (Blaauw, 1964; Brown, Dekker, and De Zeeuw, 1997). More recently, a formal distinction between (open) star clusters and associations was made by Gieles and Zwart (2011). They state that for a loosely-bound system, one dynamical time must be smaller than the age of the stellar members, whereas in unbound associations, one dynamical time exceeds the age of the stars. These systems have been further divided according to the most common spectral type of their population, or in other words how evolved their stellar population is. The three classes of association include OB, T and R, whereby their members are generally found to be OB-type stars, T Tauri stars or stars that illuminate reflection nebulae, respectively. One may therefore remark that these are the same objects seen at different evolutionary stages. Strom, Strom, and Grasdalen (1975) concluded that each group includes a range of ages across their stellar population and fundamentally cannot be used to distinguish, the cluster age.

Classically, OB associations were thought to be the expanded remnants of bound clusters that had been disrupted by some physical process (Brown et al., 1999). Residual gas expulsion has been postulated as a possible mechanism for disruption whereby stellar feedback disperses the remaining residual gas left from an initial star formation event, meaning the cluster loses mass and therefore gravitational potential energy (see Hills 1980; Lada, Margulis, and Dearborn 1984; Goodwin and Bastian 2006). Elsewhere it has been suggested that the gravitational effect of neighbouring molecular clouds can tidally heat the stellar system, causing its members to gain kinetic energy, entering a 'super-virial' state such as seen in OB associations (Spitzer Jr, 1958; Elmegreen and Hunter, 2010; Kruijssen et al., 2011). This transformation between that of a gravitationally bound system and that of an unbound stellar system is difficult to constrain observationally and requires data taken over multiple epochs in order to spatially resolve the transverse proper motions of individual clusters members. There are very few cases were this has been done (see e.g. Bouy et al. 2013; Wright et al. 2016 for the case of Cygnus OB2) but with the advent

of Gaia (Perryman et al., 2001), the proper motions of billions of galactic stars will be resolved.

### 1.3.2 The formation, evolution and survival of MSC

Stars have been observed to form out of a huge number of environments, ranging continuously in density (Longmore et al., 2014). Similarly, despite the definitions given above, stellar systems have also been observed across a continuous range of mass and stellar densities (Bressert et al., 2010). Due to a common interest in how stellar and planetary systems came to be, it is crucial to find the dependance of their nature, upon the environment from which they formed. Since the majority of stars within the universe are found to be within stellar clusters, the study of the formation of these systems is of upmost importance.

#### 1.3.2.1 Formation of young stellar systems

It is well known that stars form within molecular clouds, the primary component of which is molecular hydrogen ($H_2$). Furthermore, since the majority of stars have been found to form in clustered environments (Lada and Lada, 2003; Gutermuth et al., 2009), it is no surprise that young massive clusters are thought to form within giant molecular clouds (GMCs). These vast reservoirs of molecular gas and dust can range from $10^3$ to $10^7$ $M_\odot$ in mass, and span up to hundreds of parsecs in diameter (Murray, 2011). With reference to Table 1.1, YMCs are generally found to have a mass $> 10^4$ $M_\odot$ and thus any GMC of a similar or larger gas mass, $M_{gas}$, fulfils the most basic initial condition for YMC formation. These two parameters can be related to one another via the star formation efficiency, $\epsilon$ which can be defined within, $M_{gas} = M_\star/\epsilon$ (Longmore et al., 2014).

There currently exist two distinct formation scenarios that stem from the consideration of the initial extent of the gas reservoir, $R_{gas}$ with respect to that of the resulting cluster, $R_\star$ (Longmore et al., 2014). 'In-situ' formation describes a top-down scenario, where $R_{gas} = R_\star$. This case requires all the gas to accumulate within the volume of the resulting cluster, which is formed within one single star-formation episode. The second, bottom-up scenario, known as 'conveyor belt' formation (Longmore et al., 2014), postulates that $R_{gas}$ is several

times larger than the resulting extent of the cluster, $R_\star$. In this mechanism, the resulting cluster is formed via the merging of less-massive 'sub-clusters' that initially form in regions of much lower density compared to the 'in-situ' case. These 'sub-clusters' are thought to fall into one another and coalesce by gravitational cloud collapse and the convergence of two initially independent gas flows (Longmore et al., 2014).

Both scenarios predict a different environment prior to the onset of cluster formation and can therefore be tested observationally. Within the 'in-situ' formation scenario, prior to the single star-formation event, one should expect an object of mass, $M_{gas}$ and radius $R_\star$ with no signs of active star formation. However it is difficult to predict how long such a system would be in this phase, which depends on the ratio of the time taken to accumulate all the gas within the final cluster volume $R_\star$ and the time taken for the onset of star formation. On the other hand, the 'conveyor-belt' formation scenario suggests that one would never expect to find clouds of gas with mass $M_{gas}$ and radius $R_\star$ with no signs of active star formation (Longmore et al., 2014). In fact the observational motivation behind such a scenario is the apparent sub-structures observed in OB associations and YMCs (see Kuhn et al. 2014; Wright et al. 2016, where stars have already been formed (Banerjee and Kroupa, 2015). The observed age spreads within a particular MSC and timescale for the onset of star formation are therefore the two key observational diagnostics to differentiate between these two scenarios (Longmore et al., 2014).

The key difference between the 'in-situ' and 'conveyor belt' formation mechanisms lie within their initial conditions. The environment in which a MSC is found is therefore influential on its formation and subsequent evolution. Longmore et al. (2014) found the disk and centre of the Milky Way are utilising different formation mechanisms to assemble MSCs. In the Galactic centre, gas has been observed to reach very high densities with an absence of active star formation, implying a likely 'in-situ' cluster formation mechanism. The best example of such a cloud of gas is the infra-red dark cloud G0.253+0.016 which contains $\sim 10^5$ $M_\odot$ of gas in a radius of $\sim 3$ pc and shows almost no evidence of on-going star formation, however this has been dubbed a very extreme case (see Lis et al. 1994; Longmore et al. 2012). Conversely, gas clouds of mass $> 10^5$ $M_\odot$ typically found in the galactic disk are found to be $\sim$ 10 - 100 parsecs in size (Longmore et al., 2014),

suggesting a 'conveyor belt' formation mechanism. Furthermore, observational signatures of gravitational collapse and large scale gas flows have been found to physically 'quench' MSC progenitor clouds within the disk regions of W49 (Galván-Madrid et al., 2013) and W43 (Nguyen-Lu'o'ng et al., 2013).

#### 1.3.2.2 Evolution and survival of young stellar systems

Portegies Zwart, McMillan, and Gieles (2010) summarised the evolution of YMCs into three distinct stages. Stage 1 describes the first few Myr in which stars are still forming and a significant fraction of its mass is composed of ambient gas. After $\sim 3$ Myr when the first supernovae occur (Eggleton, 2006), the gas is dispersed and the cluster moves into stage 2, where stellar mass-loss is the largest contributor to its dynamics. Depending on the typical characteristics of the cluster, it will move into stage 3 somewhere between 100 Myr and 1 Gyr, when processes of stellar dynamics dominate its long term evolution. Stage 3 is well understood and requires relatively basic initial conditions in order to be successfully modelled by 'traditional' N-body stellar dynamical simulations. In order to simulate stage 2 however, one requires an accurate understanding of stellar evolution in order to account for stellar mass-loss and feedback, whilst also carefully considering large-scale stellar dynamics and smaller scale stellar interactions. Whilst these dynamical processes are mostly well known, our understanding of stellar evolution is far from complete. Mass-loss rates from massive stars remain an uncertainty (see e.g. Chapter 4 of this thesis and Drew 1990; Puls et al. 2006; Prinja and Massa 2010; Muijres et al. 2011), which affects our understanding of how stellar feedback influences the dynamical evolution of MSC. Furthermore, stellar feedback dominates star formation processes on scales $< 10$ pc (Kroupa, 2008) and no doubt plays a vital role in the evolution of YMCs.

The evolution of a YMC in stage 1 is composed of a complex mixture of physics, including radiative transfer, stellar evolution, gas and stellar dynamics and is currently not well understood (e.g. see Elmegreen 2007; Price and Bate 2009). This in-turn creates a large uncertainty in some of the crucial cluster properties at the outset of stage 2, the majority of which are poorly constrained by observations (Portegies Zwart, McMillan, and Gieles,

2010). Assumptions must be made regarding the stellar mass function, to what extent the cluster is mass segregated (if at all), the virial ratio, the initial spatial density and velocity distribution of the cluster members, the binary fraction and the tidal field in which the cluster resides. Each of these contribute to the subsequent evolution of the cluster.

Throughout the early evolution of YMCs (stage 1), the expulsion of the gas can occur in two ways. Gas exhaustion is a consequence of star formation, in the gravitational collapse of a molecular cloud and the following star formation process of accreting material onto a proto-star. Gas expulsion on the other hand, is driven by stellar feedback by way of proto-stellar outflows (Krumholz, Klein, and McKee, 2012), photoionising radiation (Dale, Ercolano, and Bonnell, 2012), radiative feedback (Offner, Hansen, and Krumholz, 2009) and stellar winds and supernovae (Pelupessy and Portegies Zwart, 2012). If the local gas density is high, the star formation efficiency (SFE) will be high and gas exhaustion will be the main driver of the gas depletion. However, if the local gas density (and hence SFE) is much lower, then gas expulsion will dominate the gas depletion process. In such a scenario, the gravitational potential of the system will depend on the gas (as opposed to the stars, since fewer stars have been formed; Longmore et al. 2014), which is subsequently depleted via stellar feedback, meaning the cluster will expand into a super-virial state, leaving it vulnerable to dissolution (Portegies Zwart, McMillan, and Gieles, 2010). This process is often referred to as 'infant mortality' and as a consequence of which, we observe a marked decrease in the number of young and embedded stellar clusters at the age of a few Myr (Lada and Lada, 2003). The gas expulsion phase and the importance of stellar feedback are therefore subject to the outcome of the initial star formation process and the cloud environment in which the cluster formed. Unbound, OB associations are the natural outcome of a star formation history dominated by gas expulsion whereas dense, virialised YMCs were subjected to a rich star formation history dominated by exhaustion.

A second mechanism leading to the disruption and dissolution of young stellar systems is the so called 'cruel cradle effect' (see e.g. Kruijssen et al. 2011; Elmegreen and Hunter 2000). Affecting low density clusters, tidal perturbations from dense, star forming environments disrupt and suppress the formation of bound clusters. It is worthy of note that both these disruption mechanisms depend purely upon the volume density of the young

stellar cluster.

Stellar clusters that manage to survive stage 1 have so far ejected their ambient gas. As such, star formation halts, however the systems continue to lose mass through stellar evolution. Massive OB stars on the main sequence constantly lose mass via their stellar winds before evolving off the main sequence and onto a number of possible evolutionary phases (see Section 1.1.3). If one considers a Kroupa (2001) IMF between 0.1 and 100 $M_{\odot}$, Portegies Zwart, McMillan, and Gieles (2010) state that the total cluster mass can decrease by 10%, 20% and 30% during the first 10, 100 and 500 Myr respectively.

Once a massive cluster has reached stage 3 in its lifetime, its evolution is driven by the competition between relaxation and a variety of internal heating mechanisms. Self-gravitating systems such as MSCs have no final equilibrium state (Portegies Zwart, McMillan, and Gieles, 2010). The evaporation of high velocity stars (that escape the system) and relaxation from two body-encounters ensures energy is transported from the inner to the outer regions. As a result of the core losing energy, it contracts becoming denser, whilst the outer regions expand. This also promotes mass segregation with the more massive stars sinking towards the cluster centre (Portegies Zwart, McMillan, and Gieles, 2010). Heating mechanisms within the cluster act as energy sources and can oppose core collapse. Such mechanisms include binary interactions, with 'hard binaries' being a net heat source to the cluster (see e.g. Heggie 1975; Hills 1975). Stellar collisions are also possible in the core regions where densities are high which have been cited as a mechanism to produce intermediate mass black holes (Portegies Zwart and McMillan, 2002; Gürkan, Freitag, and Rasio, 2004), further adding energy to stellar systems (Portegies Zwart, McMillan, and Gieles, 2010).

## 1.4 Cygnus OB2: a massive OB association

Founded on recent literature, this section summarises the past and presently accepted parameters of the massive stellar OB association of Cygnus OB2. In summarising the

physical characteristics of the rich stellar system, a discussion of its dynamical state will follow, including suggestions regarding the mechanisms responsible for its formation.

### 1.4.1 Location

The work presented in this thesis is centred around e-MERLIN radio observations of the Cyg OB2 association. This vast collection of stars is found within the massive star forming region of Cygnus-X that in-turn is situated in the constellation of Cygnus. Originating from the greek word for swan, Cygnus is a prominent northern hemisphere constellation, otherwise known as the Northern Cross, which lies along the plane of the Milky Way. The entire Cygnus region is home to a vast amount of star formation across many different stellar clusters and associations (see Reipurth and Schneider 2008) and has often been quoted as the most well-suited region to study the formation and evolution of such systems (see e.g. Maia, Moraux, and Joncour 2016). At relatively close proximity (0.7 - 2.5 kpc; Uyanıker et al. 2001), the Cygnus region contains 9 OB associations and at least a dozen young open clusters (Mahy et al., 2013). Figure 1.9 taken from Mahy et al. (2013) is a schematic in galactic coordinates to show the relative positions of these various stellar systems.

Cygnus-X has been identified as an immensely rich star formation region playing host to numerous stellar clusters and containing $\sim 4.7 \times 10^6$ $M_\odot$ of molecular gas (Schneider et al., 2006). It was first recorded midway through the 20th century by Piddington and Minnett (1952) as an extended radio continuum source close to Cygnus A, the 'X' denoting its extended nature. The whole region has since been well sampled by numerous surveys across a large portion of the electromagnetic spectrum (from X-rays through to the radio) with hundreds of HII regions being detected along with several supernova remnants (Uyanıker et al., 2001). X-ray observations of the region revealed the presence of a 'super-bubble' spanning 450 pc across which can be attributed to the past events of 30-100 supernova (Cash et al., 1980), at the centre of which lies the Cyg OB2 association.

With reference to the $H_\alpha$ image in Figure 1.10, 'The Great Cygnus Rift' covers the entire Cygnus regions with a variable visual extinction, $A_\nu \sim 1.0 - 6.4$ magnitudes (Dickel and Wendker, 1978, see. Section 1.4.2 for further details). Lying directly in our line of sight

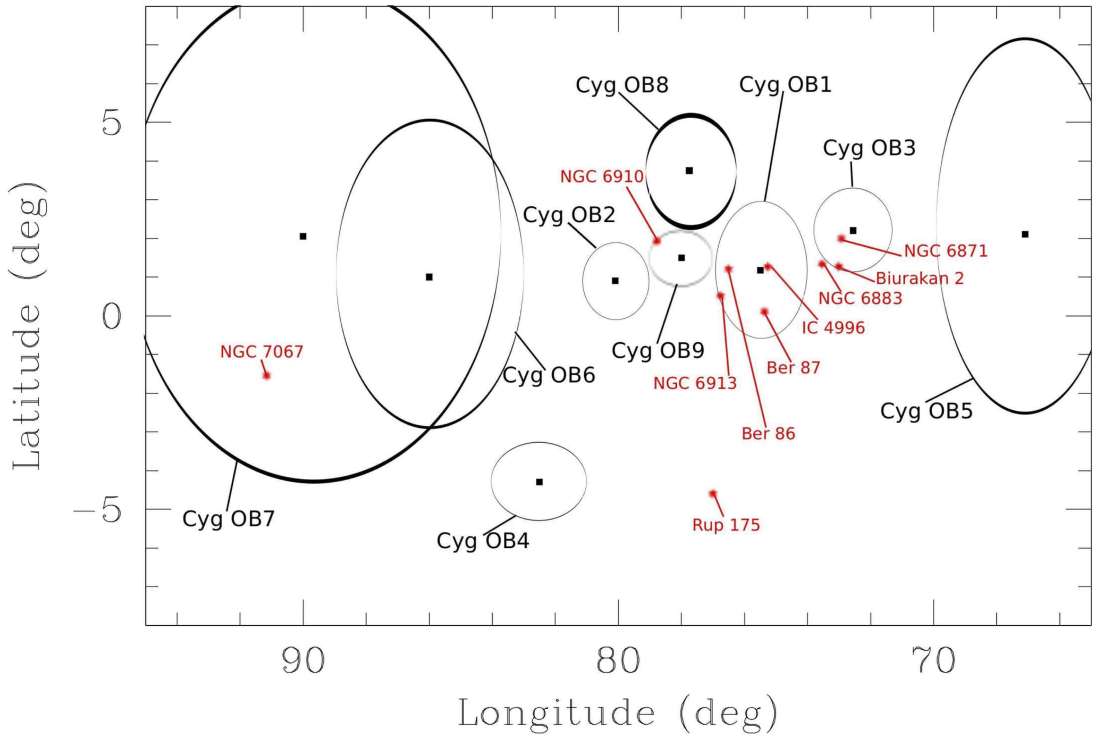



FIGURE 1.9: Relative positions of the 9 OB associations and 10 young open clusters in the Cygnus region in galactic coordinates. Image credit: Mahy et al. (2013)

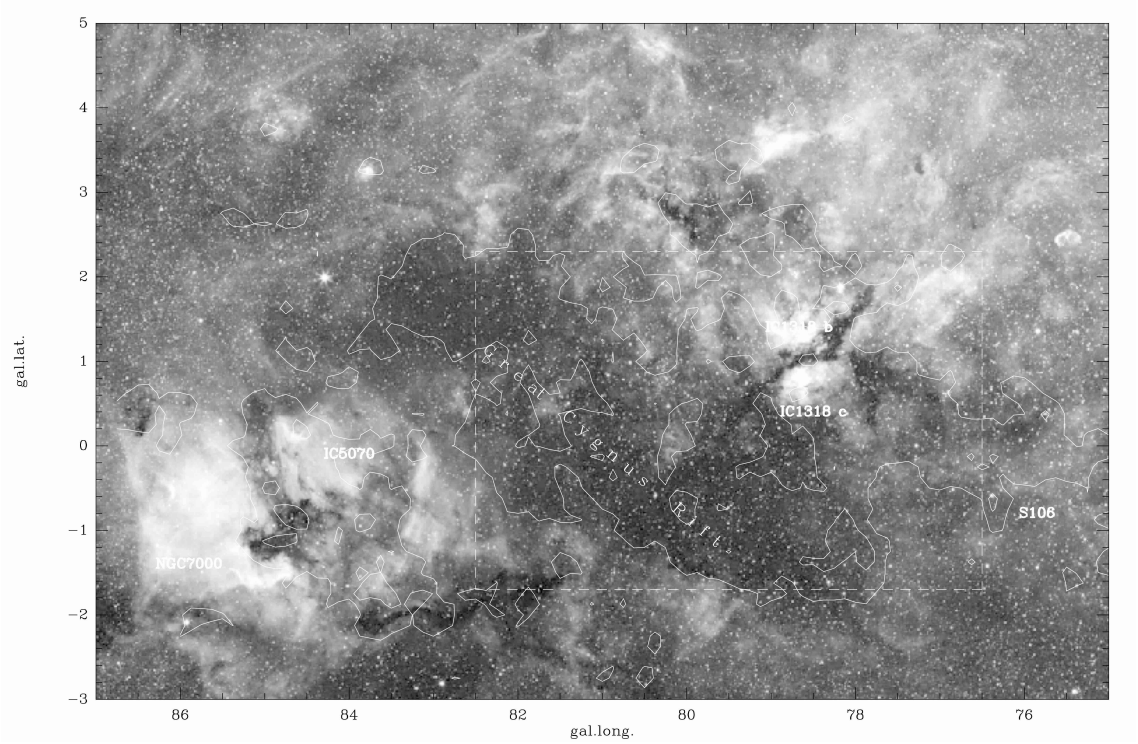



FIGURE 1.10: An $H_\alpha$ image of the Cygnus-X region in galactic co-ordinates taken from Schneider et al. (2006). The white contours of 20%, 60% and 80% of the $_{12}$CO 1→0 line emission are overlaid and were taken from the Harvard Smithsonian CfA (Centre for Astrophysics) survey by Dame et al. (1987). The large scale extinction feature is highlighted as the 'Great Cygnus Rift' which covers almost the entire region. The three well known HII regions NGC 7000 (North American Nebula), IC 5070 (Pelican Nebula) and IC 1318 are also highlighted.

towards to the Cygnus-X complex and located at a distance of $\sim$ 700 pc (Gottschalk et al., 2012), this cloud of gas and dust has typical velocities of $\sim$ 6 - 20 kms$^{-1}$ and low column densities (Schneider et al., 2007). The rift is thought to be part of the 'Gould Belt' (Schneider et al., 2007) and obscures much of the optical light from the Cygnus region.

Located in the very heart of the Cygnus-X region directly behind the 'Great Cygnus Rift' is the massive OB association of Cyg OB2. This stellar system is amongst the most massive clusters and associations found within our galaxy (Knödlseder, 2000; Hanson, 2003; Wright and Drake, 2009). First identified by Münch and Morgan (1953), its first stellar members were soon catalogued by the works of Johnson and Morgan (1954), Schulte (1956) and Schulte (1958). Since then, it has been found to harbour $\sim 2600 \pm 400$ OB stars (Knödlseder, 2000) while its total mass has been estimated to be between $10^4 - 10^5$ M$_\odot$ (Knödlseder, 2000; Hanson, 2003; Drew et al., 2008; Wright et al., 2010; Wright, Drew, and Mohr-Smith, 2015). This in conjunction with its close proximity (distance estimates span between 0.9 and 2.1 kpc, see Section 1.4.3) makes Cyg OB2 a unique case for studies of massive stellar clusters, star formation and stellar evolution within the Galaxy.

### 1.4.2 Visual extinction in Cyg OB2

Cyg OB2 suffers from a large visual extinction due to a combination of the large amounts of dust found within its parental cloud and within the 'Great Cygnus Rift' residing in the foreground. Correcting for interstellar extinction is crucial when deriving photometric and spectroscopic properties of stellar objects and if not accounted for accurately can lead to mis-classifications and erroneous basic parameters. Figure 1.11 shows an extinction map of the entire Cygnus-X region, centred on Cyg OB2 and obtained using the Two Micron All Sky Survey (2MASS; Cutri et al. 2003) *JHK* infrared observations. The dark colours represent a higher visual extinction, Cyg OB2 is highlighted in the centre of the image against dark patches of interstellar extinction.

The pioneering work of Johnson and Morgan (1954) first identified 14 members of the association, reporting on a high extinction towards these stars with a reddening parameter E($B-V$) ranging from 1.4 to 2.3. Furthermore, the association member and LBV candidate

Cyg OB2 #12 was once known as the most reddened star known at optical wavelengths with a value of E($B-V$) = 3.2 (Souza and Lutz, 1980). Since the work by Johnson and Morgan (1954), numerous photometric studies (Schulte, 1958; Reddish, Lawrence, and Pratt, 1967; Torres-Dodgen, Carroll, and Tapia, 1991; Massey and Thompson, 1991; Hanson, 2003; Drew et al., 2008; Wright, Drew, and Mohr-Smith, 2015) have derived the extinction properties for a large variety of Cyg OB2 cluster members. In the most recent study, Wright, Drew, and Mohr-Smith (2015) derive the individual extinction to 164 primary OB stars within the association with values ranging from $A_V = 2.2^m$ to $10.2^m$ (where the superscript $m$ denotes magnitude). They first derive the extinction law towards the association using the unreddened colours of Martins and Plez (2006) for the O-type stars and those from Fitzgerald (1970) and Koornneef (1983) for the B-type stars. Selecting a high-confidence subset of 18 single O-type stars (spectral types taken from Massey and Thompson 1991), Wright, Drew, and Mohr-Smith (2015) fit their spectral energy distribution (SED) using reddened stellar spectra calculated using the non-LTE atmosphere code CMFGEN (Hillier and Miller, 1998) to infer the reddening coefficients. In combination with the unreddened O-star colours of Martins and Plez (2006) they obtained the reddened photometry for these stars as a function of spectral type, $R_V$ and $A_V$. Furthermore, the spectral classifications of their sample of 164 primary OB stars were obtained from the literature, where multiple classifications were found, they adopted those derived from the highest spectral resolution. Utilising their derived extinction law, and a ratio of total to selective extinction $R_V = 2.91$ as calculated by Fitzpatrick and Massa (2007), they derived the extinction towards each of the 164 OB members, finding a median of $A_V = 5.4$ magnitudes.

### 1.4.3 Distance to Cyg OB2

The distance to the Cyg OB2 association remains an uncertainty, with estimates in the literature spanning between 0.9 kpc and 2.1 kpc. Using both spectroscopic and photometric observations to infer its distance becomes complicated since the region is known to suffer from variable extinction (see Section 1.4.2). Its location within the Galaxy also adds further complication since it lies at $l = 80°$ where the relation between radial velocity and distance

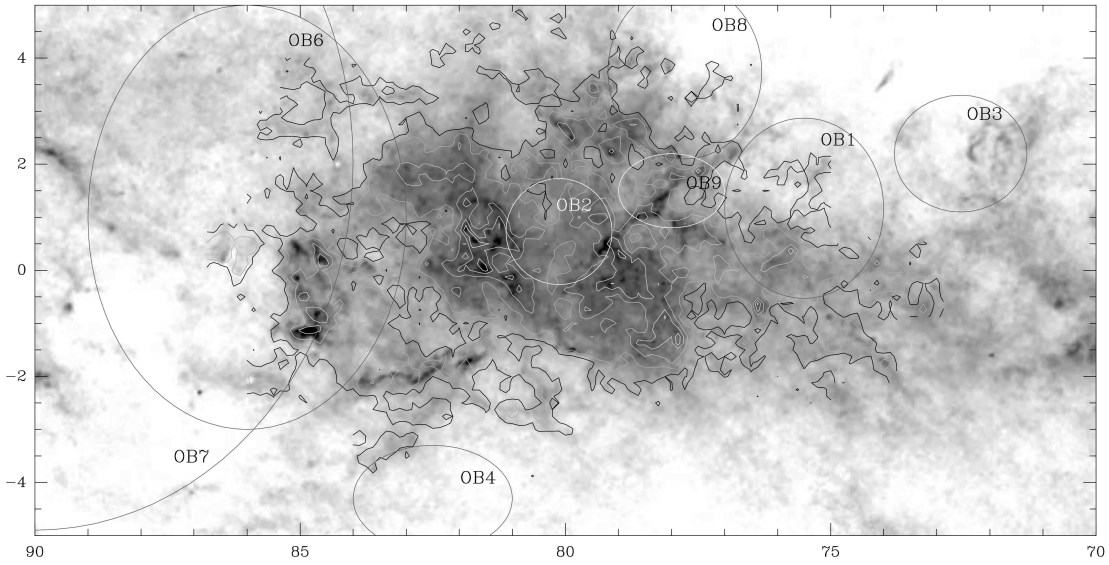



FIGURE 1.11: A visual extinction map in galactic coordinates of the Cygnus X region (image credit Reipurth and Schneider 2008). The greyscale image represents $A_\nu$ spanning in magnitudes from 1 (white) to 32 (black) and was obtained from near-infrared 2MASS data. The contours represent the line integrated (from -10 to 20 $kms^{-1}$) emission of the $_{12}CO$ 1→0 transition from 20% to 100% of the maximum intensity of 39.7 $kms^{-1}$ in steps of 20%. Overlaid, are ellipses that represent the approximate extent of the OB associations within the region as in Figure 1.9 taken from Uyanıker et al. (2001).

is poorly defined (see. Dame and Thaddeus, 1985; Dame, Hartmann, and Thaddeus, 2001). Massey and Thompson (1991) used spectroscopic and photometric observations of 63 stars in Cyg OB2 to infer a distance of 1.71 kpc to the association, a value in good agreement with previous studies (Humphreys, 1978; Torres-Dodgen, Carroll, and Tapia, 1991). The MK optical spectra of 14 Cyg OB2 stars were used by Hanson (2003) to derive a much closer distance estimate (D $\sim$ 1.45 kpc). Subsequent work by Negueruela et al. (2008) supported this revised distance in fitting model isochrones to a semi-observational HR diagram of Cyg OB2. Taking a different approach, Linder et al. (2009) analysed the light curve of the eclipsing binary Cyg OB2 #5, yielding a distance of 0.90−0.95 kpc. Whilst reporting that their result is in need of confirmation, they highlight the implications this distance estimate would have upon the stars within Cyg OB2, greatly reducing both the luminosities and mass-loss rate of its massive star members.

More recently, measurements of the trigonomic parallaxes and proper motions of five star-forming regions within the Cygnus X complex gave a distance of 1.40 $\pm$ 0.08 kpc to the

region (Rygl et al., 2012). Furthermore Kiminki et al. (2015) measured the distance to four eclipsing binary members of Cyg OB2, obtaining a weighted average distance of 1.33 ± 0.06 kpc. This result sits slightly lower than those obtained via spectro-photometric methods, although interestingly Hanson (2003) noted that if they were to adopt the cooler $T_{eff}$ scales of Martins, Schaerer, and Hillier (2002) and Martins, Schaerer, and Hillier (2005), their distance estimate would reduce to 1.2 kpc. This was rejected by Hanson (2003) as moving the association closer would reduce the luminosity of the association's supergiants and hence also reduce their mass-loss rate, making them more discrepant than those predicted from stellar wind theory (e.g. Vink, Koter, and Lamers, 2000). Growing evidence exists for the need to revise accepted mass-loss rates downwards (see. Puls, Vink, and Najarro, 2008a) and indeed this uncertainty is much of the motivation behind this thesis. For the remainder of this thesis and within the subsequent scientific analysis that follows, a distance to the Cygnus OB2 association of 1.4 ± 0.1 kpc has been adopted.

### 1.4.4 Cluster members

Since the first detection of 'blue giant' stars in the region of Cyg OB2 (Münch and Morgan, 1953), numerous observations, across the entire electromagnetic spectrum have revealed a rich and diverse population of stellar objects (see Table 1.2). The *UBV* photometric and spectroscopic observations by Massey and Thompson (1991) revealed a wealth of OB stars before the work of Knödlseder (2000) utilised *JHK* photometric observations from 2MASS to estimate the association to contain ∼ 2600 ± 400 OB stars. More recently, the recent census compiled in Wright, Drew, and Mohr-Smith (2015) list 166 OB stars with spectral classification, including 52 O-type stars. Kobulnicky et al. (2014) present results from the Cyg OB2 Radial Velocity Survey that took place between 2001 and 2011. Their spectroscopic observations reveal the orbital parameters of 48 massive multiple-star systems within Cyg OB2, concluding a binary fraction near 55% for systems with a period less than 5000 days. At much shorter wavelengths, X-ray observations of Cyg OB2 have been compiled as part of the Chandra Cygnus OB2 Legacy Survey (Wright et al., 2014b). The central 0.5 deg$^2$ of the association has been found to contain ∼ 8000 X-ray point sources, with many of them considered to be Young Stellar Objects (YSOs) and T Tauri

stars (or at least 444 have been found to be stars with disks; Guarcello et al. 2015). In addition to the vast number of YSOs, and massive OB-type stars, the association is home to a number of Be stars, two Wolf-Rayet stars (WR145, WR 146), two candidate Luminous Blue Variable (LBV) stars (G79.29+0.46, Cyg OB2 #12), a red supergiant (IRC+40 427), a B[e] star (MWC 349) and a very high energy $\gamma$-ray source (TeV J2032+4130).

TABLE 1.2: To highlight the main observations and surveys that have been centred on or that have included members of the Cygnus OB2 association.

| **Observation Type** | **Reference** |
|---|---|
| Optical Photometry & Spectroscopy | Johnson and Morgan (1954) |
| $UBV$ Photometry & Optical Spectroscopy | Schulte (1958) |
| $UBV$ Photometry | Lawrence and Reddish (1965) |
| Optical Spectroscopy | Walborn (1973) |
| $JHKL$ Photometry | Voelcker (1975) |
| X-Ray Observations | Harnden et al. (1979) |
| X-Ray Observations | Cash et al. (1980) |
| $\lambda$6cm Radio Observations (VLA) | Abbott et al. (1980) |
| Optical Spectroscopy | Hutchings (1981) |
| $HKLMN$ Photometry & Spectroscopy | Leitherer et al. (1982) |
| Radio Observations | Persi et al. (1985) and Persi et al. (1985) |
| Radio Observations | Bieging, Abbott, and Churchwell (1989) |
| $JHKL$ Photometry | Torres-Dodgen, Carroll, and Tapia (1991) |
| Optical Photometry & Spectroscopy | Massey and Thompson (1991) |
| Radio (VLA) Observations | Condon et al. (1998) |
| VRI Photometry | Pigulski and Kołaczkowski (1998) |
| $JHK$ (2MASS) Photometry | Knödlseder (2000) |
| $JKH$ Photometry | Comerón and Torra (2001) |
| near-IR ($KH$-band) Spectroscopy | Comerón et al. (2002) |
| IPHAS Survey | Drew et al. (2005) |
| UKIRT ($JHK$) Infrared Photometry | Hewett et al. (2006) |
| X-Ray and near-IR Observations | Colombo et al. (2007a) |
| Radio Observations (VLA/GMRT) | Martí et al. (2007) |
| UKIDSS ($JHK$) Infrared Photometry | Lucas et al. (2008) |
| Optical Spectroscopy & Photometry | Vink et al. (2008) |
| X-Ray (Chandra) Observations | Wright and Drake (2009) |
| Spitzer/IRAC Legacy Survey (3.6, 4.5, 5.8, 8.0, 24$\mu$m) | Beerer et al. (2010) |
| SDSS Catalogue (8th Release) | Aihara et al. (2011) |
| HST Fine Guidance Sensor Survey | Caballero-Nieves et al. (2011) |
| Optical Spectroscopy | Comerón and Pasquali (2012) |
| $RIZ$ Optical Photometry | Guarcello et al. (2012) |
| Optical Spectroscopy | Kobulnicky et al. (2014) |
| X-Ray Observations | Wright et al. (2014b) |

TABLE 1.3: The fundamental physical parameters that describe the Cygnus OB2 association. The RA and DEC describe the centre of the association as chosen by numerous authors due to the position of the main concentration of OB stars.

| | | Reference |
|---|---|---|
| RA (J2000) | $20^h33^m12^s$ | 1 |
| DEC (J2000) | $+41°19'00"$ | 1 |
| Stellar Mass | $16500^{+3800}_{-2800}$ $M_\odot$ | 2 |
| Virial Mass | $(9.3 \pm 0.8) \times 10^5$ $M_\odot$ | 3 |
| Volume Density | $\sim$ 100 stars $pc^{-3}$ | 4 |
| Age | 1-7 Myr (Peak 4-5 Myrs) | 2 |
| Distance | $1.4 \pm 0.08$ | 5 |
| Visual Extinction, $A_V$ | $4^m - 7^m$ (IQR) | 2 |
| OB Members | $2600 \pm 400$ | 6 |
| Binary Fraction | 55% | 7 |
| 3D Velocity Dispersion | $17.8 \pm 0.6$ km $s^{-1}$ | 3 |
| Half-Light (Effective) Radius | $10.1 \pm 0.9$' (4.1pc) | 3 |
| Observational Diameter | $\sim 2°$ ($\sim$ 49 pc) | 6 |

References: **1** Wright et al., 2014b, **2** Wright, Drew, and Mohr-Smith, 2015, **3** Wright et al., 2016, **4** Wright et al., 2014a, **5** Rygl et al., 2012, **6** Knödlseder, 2000, **7** Kobulnicky et al., 2014.

### 1.4.5 General properties

Some of the typical cluster characteristics of the Cygnus OB2 association are summarised in Table 1.3, including an appropriate reference. The centre of the association, as defined in Table 1.3, was chosen as the centre of the 36 pointing array defined by the observations made in the Chandra Cygnus OB2 Legacy Survey (Wright et al., 2014b) and corresponds to the main concentration of OB stars within the system. The stellar mass is taken from the recent census of OB-type stars within the association by Wright, Drew, and Mohr-Smith (2015) who gathered and homogenised data from across the literature. In deriving the mass function, they found the total mass of the system to be $16500^{+3800}_{-2800}$ $M_\odot$ , a value slightly lower than previous estimates such as $(4 - 10) \times 10^4$ $M_\odot$ from Knödlseder (2000). Furthermore, in fitting stellar evolution models (including rotation) from Ekström et al. (2012), they find a large spread in the age of Cyg OB2, concluding that a continuous bout of star formation occurred between 1 and 7 Myr that peaks around 4-5 Myr. This result is consistent with studies of the association's low and intermediate mass members (e.g. Drew et al. 2008; Wright et al. 2010).

The volume density and physical structure of Cyg OB2 has been investigated using a

Chandra X-ray selected sample of young stars from the observations of Wright and Drake (2009). This was explored by Wright et al. (2014a) who found the association to be of low volume density ($\sim$ 100 stars pc$^{-3}$) yet exhibiting considerable physical structure. They also found no evidence of mass segregation, concluding that at present, Cyg OB2 is close to its initial conditions, showing no signs of dynamical evolution. This picture is inconsistent with the model that all stars form in highly dense clusters and instead suggests that stars up to a mass of $\sim$ 100 M$_\odot$ (as have been found in Cyg OB2) can indeed form in low density environments. In the particular case of massive star formation, these findings suggest that massive stars form as scaled up versions of low-mass star formation (e.g. Shu, Adams, and Lizano 1987; McKee and Tan 2003) and argue against competitive accretion theory (Bonnell, Vine, and Bate, 2004) and through merging events (Zinnecker and Yorke, 2007) by which massive stars form.

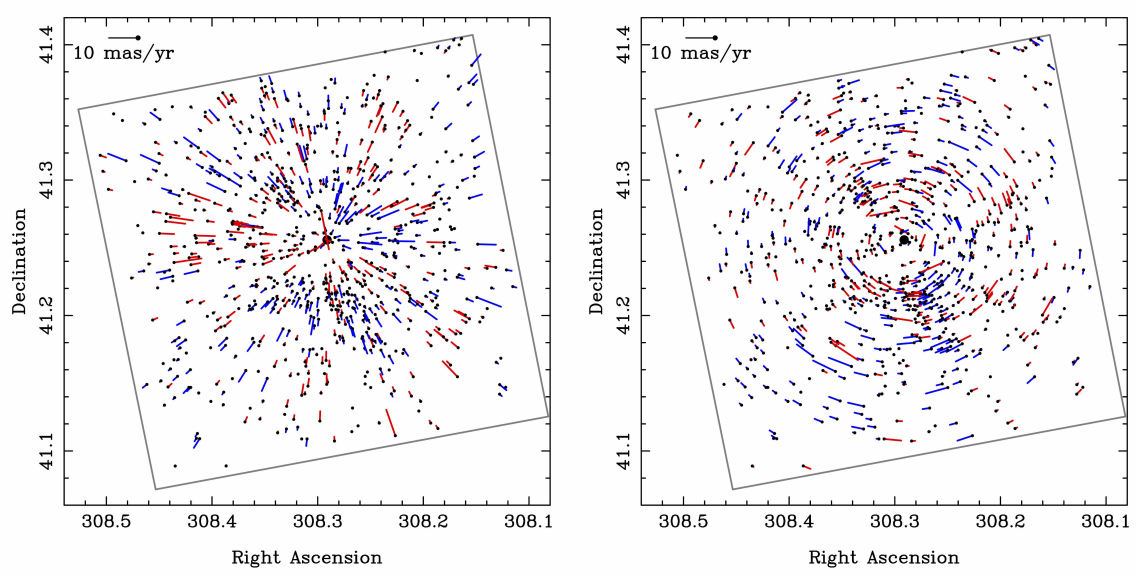



FIGURE 1.12: Results of the high-precision proper motion (PM) study of 798 stellar members of Cyg OB2, as conducted by Wright et al. (2016) as part of the DANCe project. The left panel shows the radial component of the PM vectors, with blue representing stars moving outwards from the cluster centre and red representing those moving towards the centre. The right panel shows the transverse component of the PM vectors, blue representing stars moving clockwise, red - anti-clockwise. The black dots denote the nominal position of the association centre.

More recently, a high-precision proper motion study of Cyg OB2 was conducted as part of the Dynamical Analysis of Nearby Clusters (DANCe) project by Wright et al. (2016). In measuring the proper motion of 873 X-ray and spectroscopically selected stars, they derive

a 3-dimensional velocity dispersion of $17.8 \pm 0.6$ km s$^{-1}$ and a virial mass significantly larger than the observed stellar mass (see Table 1.3). This is a direct confirmation that the association is gravitationally unbound. The observed kinematical sub-structures and non-isotropic velocity dispersions again suggest a lack of dynamical evolution, implying Cyg OB2 was likely born highly sub-structured and gravitationally unbound. Furthermore, Wright et al. (2016) find no global expansion pattern in the proper motions of the associations members (see Figure 1.12). This argues against the model of residual gas expulsion or tidal heating causing the dissolution of the system and the classical view that OB associations are expanded remnants of disrupted star clusters (Brown et al., 1999; Lada and Lada, 2003). Despite the fact that the association has already expelled the majority of its initial gas reservoir (Wright et al., 2016), there remains no signature of gas expulsion on the kinematic properties of Cyg OB2 and certainly no evidence that gas expulsion is responsible for its global unbinding. The local substructures are found to be in (or close to) virial equilibrium, giving further evidence of the little impact any gas expulsion has had. These findings therefore support the theoretical work of Kruijssen et al. (2012) and favour the scenario of gas exhaustion (see Section 1.3.2.2), whereby the initial gas reservoir is depleted by its accretion onto stars themselves which gives further support to the observations of very high mass stars found within Cyg OB2 and the massive star formation scenario of Shu, Adams, and Lizano (1987) and McKee and Tan (2003). This places Cyg OB2 as an outlier, since the majority of massive stellar systems found within the low density environment of the spiral arms of the Milky Way favour gas expulsion as the dominant gas depletion mechanism (see Section 1.3.2.2). Wright et al. (2016) conclude from their PM study, that Cyg OB2 was not born as a single dense cluster but was instead born with a considerable amount of physical and kinematic sub-structure that is currently observed to be close to, or in virial equilibrium.

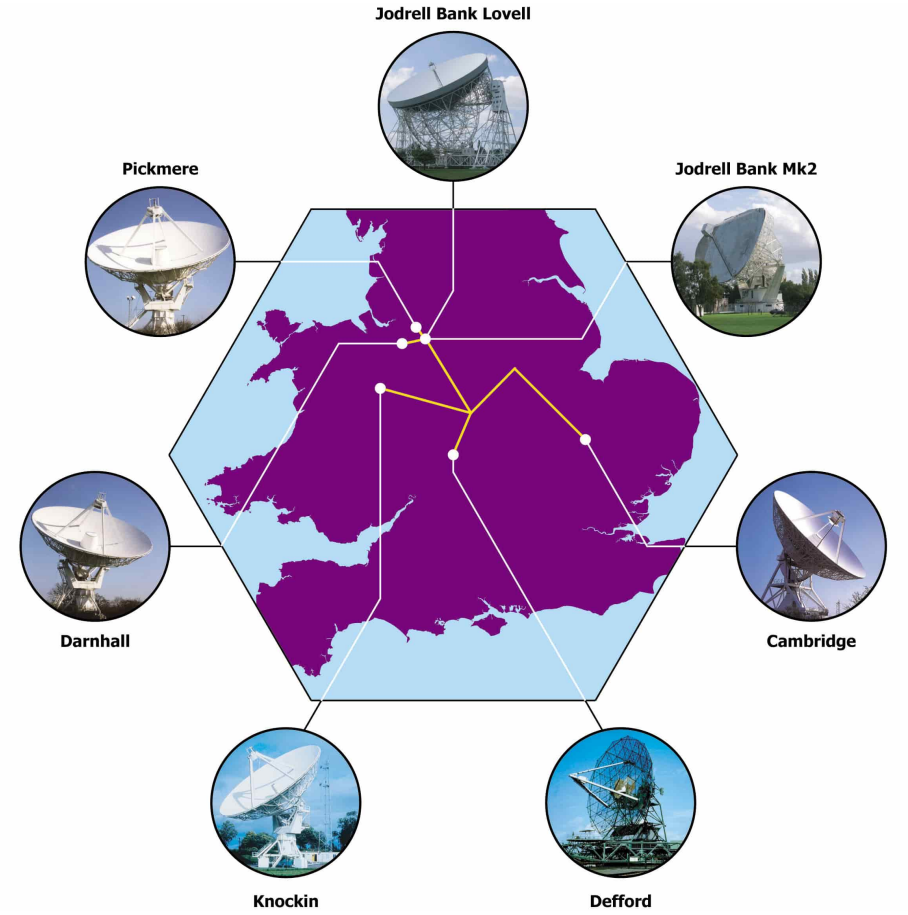



FIGURE 1.13: The positions of the seven antennas across the United Kingdom that make up the e-MERLIN radio interferometer.

## 1.5 The e-MERLIN array

e-MERLIN is an upgrade to MERLIN (Multi-Element Radio Linked Interferometer Network) operated solely by the University of Manchester. Funded by the Science and Technology Facilities Council (STFC) it consists of seven radio telescopes all within England (Figure 1.13 shows the location of each antenna). There are currently three observing bands for e-MERLIN, L-band (1.3-1.8 GHz), C-band (4-8GHz) and K-band (22-24GHz). The upgrade consisted of a completely new optical fibre network between the antennas, a new WIDAR (Wideband Interferometric Digital ARchitecture) correlator based at Jodrell Bank and new receivers, increasing the bandwidth by approximately 2 orders of magnitude. This in turns significantly increases the $u,v$ coverage of e-MERLIN, compared to that of its predecessor, resulting in a continuum sensitivity increase of a factor of $\sim 10$.

As an interferometer it is fairly unique with a longest baseline of 217 km between the Lovell and Cambridge antennas and the shortest useable baseline of Mark II to Pickmere antennas at 11 km. The smallest baseline is 400m and connects the Lovell and Mk2 antennas, however the lack of similar length baselines creates a large gap in the *uv*-plane meaning large spatial scales on the sky are subsequently under-sampled; as such this baseline is omitted from use. The length of e-MERLIN baselines puts the array in an interesting position in regards to other interferometers. Its longest baseline surpasses that of most interferometric arrays, yet falls very short of the baseline length used by the VLBA (Very Long Baseline Array) network. As such e-MERLIN is sensitive to a unique set of spatial scales on the sky, capable of reaching resolutions of 150, 40, 12 mas with sensitivities between $5-6$, $1.8-2.3$ and $\sim 15$ $\mu$Jy/bm for the L-, C- and K-bands respectively. For the full technical aspects of the e-MERLIN array, please refer to Table 1.4. These sensitivities at L- and C-band are precisely what COBRaS aims to take advantage of in the detection of radio flux from massive stars within the Cyg OB2 association.

TABLE 1.4: The basic observing capabilities of the e-MERLIN array.

| | L-band (21cm) | C-band (6cm) | K-band (1.3cm) |
|---|---|---|---|
| Resolution (mas) | 150 | 40 | 12 |
| Field of View (arcmin) | ∼11.4 | ∼3.2 | ∼0.7 |
| Freq. Range (GHz) | 1.3-1.7 | 4-8 | 22-24 |
| Bandwidth (GHz) | 0.4 | 2.0 | 2.0 |
| Sensitivity ($\mu$Jy/beam) | 5-6 | 1.8-2.3 | ∼15 |
| Surface Brightness Sensitivity (K) | ∼190 | ∼70 | ∼530 |
| Astrometric Performance (mas) | ∼2 | ∼1 | ∼2 |
| Amplitude Calibration | 2% | 1% | 10% |

All the values presented here assume the presence of the 76m Lovell telescope in the array. The resolution has been calculated using a uniform weighting at the central frequency. Furthermore the sensitivity and surface brightness temperature have been calculated assuming a full imaging run ($\sim$ 12 hours on source). The astrometric performance quoted from a typical 3 degree target-calibrator separation using the Very Long Baseline Array (VLBA) calibrator survey and the amplitude calibration applies to targets for day-to-day repeatability. This information was taken from the e-MERLIN webpage: Beswick (2013).

## 1.6 COBRaS - an e-MERLIN Legacy survey

The Cygnus OB2 Radio Survey (COBRaS) is an e-MERLIN Legacy project (P.I. Prof Raman Prinja) that has the goal of assembling a comprehensive and long lasting radio

dataset of the Cygnus OB2 association. Designed to exploit the enhanced capabilities of e-MERLIN, the project is conducting a highly sensitive, targeted, deep-field mapping of the most massive OB association known within our Galaxy. In the delivery of the most up-to-date and detailed radio census of the tremendously OB rich stellar system, the survey will ultimately offer a direct comparison to open clusters, young massive clusters and OB associations in the Galaxy, whilst offering a unique insight into the type of environments found in the larger 'super star' clusters found in other galaxies. COBRaS aims to advance our understanding of current astrophysical problems within two related, key areas of research:

1. The mass-loss, stellar winds and evolution of massive stars.
2. The massive binary frequency and the incidence of non-thermal radiation.

The specific goal of the project is explained in further detail in Section 1.6.1 and will provide the motivation behind Chapters 4 and 5 of this thesis. The COBRaS project was awarded a total allocation of 252 hours and 42 hours at C- (6 cm) and L-band (21 cm) respectively. The substantial volume of data will herald crucial results to our understanding of massive stars and their role in YMCs. Furthermore, in combination with current multi-waveband surveys of the Cyg OB2 association (IPHAS: H$\alpha$, Spitzer: near-IR, Chandra: X-ray) and indeed future surveys (e.g. JWST, Gaia), COBRaS will deliver a dataset of lasting value that will also steer additional archival studies in both stellar and extra-galactic astrophysics.

### 1.6.1 Scientific aims

#### 1.6.1.1 Mass-loss, stellar winds, and massive star evolution

As explored earlier throughout Section 1.2, there currently exist substantial uncertainties in the mass-loss of massive stars. This fundamental process, responsible for the recycling of energy and material back into the ISM is in doubt. The mass-loss rates of luminous OB stars inferred from different observational diagnostics have been found to be discordant,

challenging the accepted model of mass-loss from stellar winds. Results from these diagnostics, namely observations of the H$\alpha$ line ($\lambda$ =656nm), free-free emission in the sub-mm and radio, and those from ultraviolet (UV) resonance lines, have been found to differ on the order of magnitude scale (Prinja, Massa, and Searle, 2005; Fullerton, Massa, and Prinja, 2006; Puls et al., 2006). This discrepancy has far reaching consequences affecting the evolution and eventual fate of massive stars, the chemical enrichment of the local environment and that of the Galaxy, and the amount of material and energy available for the future generations of stars.

Massive stars have a huge influence in many areas of astrophysics. In the injection of energy, momentum and material back into the ISM, they drive numerous physical processes including turbulence, ionisation, heating and chemical mixing, all of which play a large role in planet formation and eventually the building blocks of life. Massive stars generate a huge amount of UV radiation, are the sites of cosmic ray acceleration and severely shape the environment in which they reside. Whilst creating kilo parsec scale super bubbles and wind-wind collision regions, they are also of great significance regarding the evolution of the universe as contributors to the epoch of re-ionisation. Their importance outlines the need for an accurate understanding of their fundamental parameters such as the extent of their mass-loss via clumped and/or porous radiation driven winds.

Mass-loss is a fundamental parameter that is key to the accuracy and success of stellar evolution models. The amount of mass shed by a massive star per unit time dictates its time spent on a particular evolutionary sequence and furthermore, promotes the existence of Wolf-Rayet (WR) stars, neutron stars, supernovae and black-holes by driving stellar evolution (see Section 1.1.3). OB stars have been found to lose mass via their radiation driven stellar winds (see Section 1.2.1 for further details). The electron-ion interactions that occur within their ionised winds produce radiation at radio wavelengths through thermal free-free emission. As shown in Section 4.2.2, this radio emission can be used to infer the mass-loss rate. As a $\rho^2$ process, it is very sensitive to any wind-structure (or clumps) that may be present within the wind. Radio inferred mass-loss rates in comparison to the H$\alpha$ or UV line diagnostics, potentially offer the most straightforward and least model dependant determination. Since the radio emission arises at large stellar radii from the photosphere

where the terminal velocity of the wind will have been reached, this technique requires no knowledge of the photospheric profile or the velocity law of the wind. Furthermore, the radio flux is not strongly dependant on the ionisation conditions of the wind but along with observations in other spectral regions (mm, near-IR, H$\alpha$), can be used to constrain the run of the clumping factor as a function of stellar radii (Puls et al. 2006; see also Chapter 4 for further details).

In utilising e-MERLIN's enhanced sensitivity to survey the $\sim 0.5$ deg$^2$ area of the Cygnus OB2 association at both L- and C-band, COBRaS aims to detect the thermal free-free radio emission, from a large number of the estimated 2600 $\pm$ 400 OB stars within the association. This will allow for reliable mass-loss estimates for a huge number of massive stars across a variety of evolutionary states. A comparison with other datasets across multiple wavebands will allow for an investigation into the amount of clumping within OB star winds. Moreover, COBRaS will conduct a study of the radio region clumping factor as a function of spectral type (and therefore as a function of stellar luminosity, effective temperature, etc), allowing for powerful constraints to be made on physical models for the origins of clumping.

#### 1.6.1.2 Binarity and the incidence of non-thermal radiation.

The detection of non-thermal radiation in a cluster such as Cyg OB2 is highly indicative of a massive star binary since its production in single massive stars has been shown to be unlikely (van Loo et al. 2008; see also Chapter 5). In, for example, an O+O type binary, the respective stellar winds collide causing shocks, around which electrons are accelerated to relativistic velocities. These relativistic electrons emit synchrotron radiation that can be detected at radio wavelengths (Dougherty and Williams, 2000). A detected source from a single-epoch radio observation that is shown to be non-thermal and prior knowledge (from existing catalogues) that the object is indeed a massive star, is all that is required to identify massive star binaries. With observations at both L- and C-band, COBRaS will accurately determine the spectral index of each source detected. Cross-correlation of the results with existing catalogues will infer a better determined binary frequency in Cyg

OB2, a hugely important parameter for both evolutionary population synthesis models and dynamical simulations of stellar clusters. This in turn will have a broad impact on our understanding of the evolution and survival of massive star clusters and the chemical evolution of the Galaxy. Furthermore, massive star binaries have become an explanation for the existence of observed ratios between various post-main sequence objects. A better estimate of the binary frequency in massive stellar clusters will undoubtedly highly benefit a wide range of astrophysical science.

The study of non-thermal emission from the colliding wind region of massive star binaries is currently limited to a handful of examples, with three of these systems residing in the Cyg OB2 association (Contreras et al., 1997; Rauw et al., 2002; Blomme et al., 2005; van Loo et al., 2008; Blomme et al., 2010). As such, the massive star binaries detected with COBRaS will allow for a statistical investigation into the dependance of the non-thermal emission upon the binary parameters. Moreover, in the case of short period binaries, the detectability of synchrotron emission is uncertain due to the considerable amount of free-free absorption within their stellar winds (see Dougherty and Williams 2000). Cyg OB2 #8A, a massive star binary with a period of just 21.908 $\pm$ 0.040 days (Blomme et al., 2010) has been shown to display phase-locked variations in its radio flux. The authors Blomme et al. (2010) model its behaviour to find that the synchrotron emitting region is not completely hidden by the free-free absorption. They also indicate that the clumping or porosity in the stellar wind could explain their observed radio spectral index and hence that non-thermal emitters such as Cyg OB2 #8A could also be used to investigate small-scale wind structure. Whilst the detection of massive star binaries will give us a better estimate of the binary frequency, they will also allow for the study of the colliding wind region itself in relation to the binary parameters and indeed the Fermi mechanism thought responsible for accelerating electrons to relativistic speeds.

Non-thermal radio emission is not only produced in the colliding wind regions of massive stars, but also from a variety of other pre-main sequence (PMS) objects such as Young Stellar Objects (YSOs) and T Tauri stars. The radio emission is produced due to the large amount of magnetic activity present in these PMS objects. Using the pre-existing X-ray catalogues and the relation between the X-ray and Radio luminosities (Güdel et

al., 2002), there appears a substantial number of sources within Cyg OB2 that should be detectable with the COBRaS observations. Furthermore, flaring events will only increase the detectability of these objects. T Tauri stars are divided depending on the equivalent width of their H$\alpha$ line into Classical T Tauri stars (CTTS; H$\alpha$ width $>$ 10 Å) and Weak T Tauri Stars (WTTS; H$\alpha$ $<$ 10 Å). These two classes of object will be able to be inferred since WTTS are found to be non-thermal emitters and CTTS appear to have similar radio properties to the PMO Herbig Ae/B stars, which are predominantly thermal in nature. The recent Chandra X-ray survey of Wright et al. (2014b) also determined $\sim$ 8000 X-ray point sources within the central 0.5 deg$^2$ of Cyg OB2 with many of them considered to be YSOs. In cross-correlating the COBRaS radio catalogue with their X-ray findings, one will be able to directly test the relationship between the radio and X-ray luminosities for a large variety of YSOs.

### 1.6.2 Technical aspects

The observations obtained with the e-MERLIN interferometer (see Section 1.5) must fulfil some basic requirements in order for COBRaS to meet the scientific goals outlined above. The few OB stars that have been previously detected at radio wavelengths in Cyg OB2 have flux densities of the order of a few mJy. The planned observations with e-MERLIN at L- (21cm) and C-band (6cm) aim to reach sensitivities of $\sim$ 10 $\mu$Jy and $\sim$ 3 $\mu$Jy respectively. This would ensure a respective 5$\sigma$ detection threshold of 50 and 15 $\mu$Jy at L- and C-band, allowing for the detection of numerous OB stars, YSOs and radio bright extragalactic objects.

#### 1.6.2.1 Observation strategy

In order to cover a large enough area of the Cygnus OB2 association, both observing bands require multiple (over-lapping) pointings. Each observation strategy can be seen in Figure 1.14, where seven pointings are required to sufficiently sample the association at L-band, whilst the smaller primary beam of e-MERLIN at C-band requires 42 over-lapping pointings in order to sample a similar area. To meet the required sensitivities, each L- and

C-band pointing needs $\sim$ 5 and 6 hours of on-source integration time respectively. In order to facilitate a good $u,v$ coverage, the observations will be cycled into 2 $\times$ 7 minute scans per pointing. Over the course of each observing run, each pointing will therefore have the same $u,v$ coverage.

#### 1.6.2.2 Data treatment

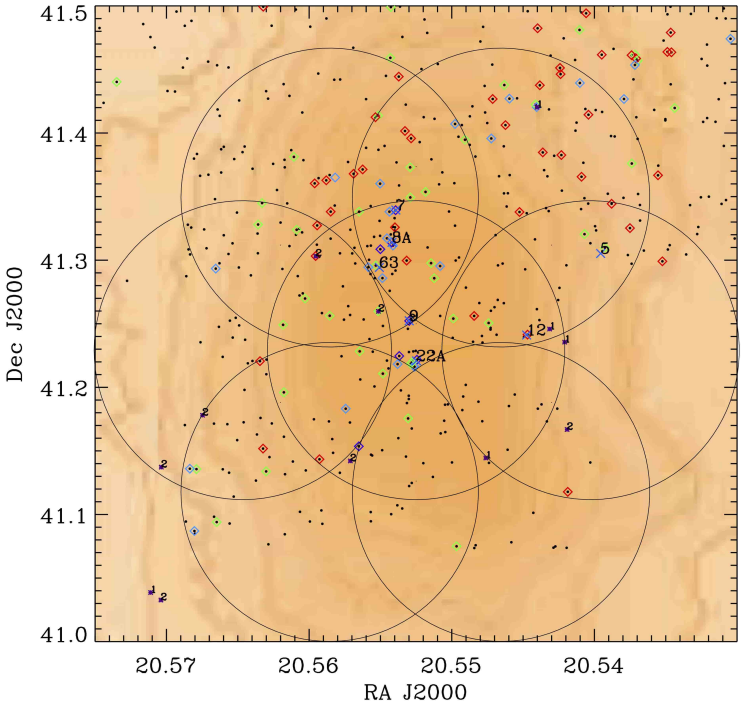
A thorough explanation of the data treatment in regards to the L-band dataset can be found in Chapter 2. A few of the key processes and steps involved from obtaining the raw data, through to its scientific analysis from a technical prospective are outlined below.

1. *Data Processing* - the large volume of data ($\gtrsim$ 1 TB) will require adequate scripting procedures in order to automate as many processes as possible throughout the entire data treatment. The Astronomical Imaging Processing System (AIPS) will be used to treat, calibrate and image the COBRaS with scripting procedures written in ParselTongue, a Python interface to AIPS (see Chapter 2 for further details).

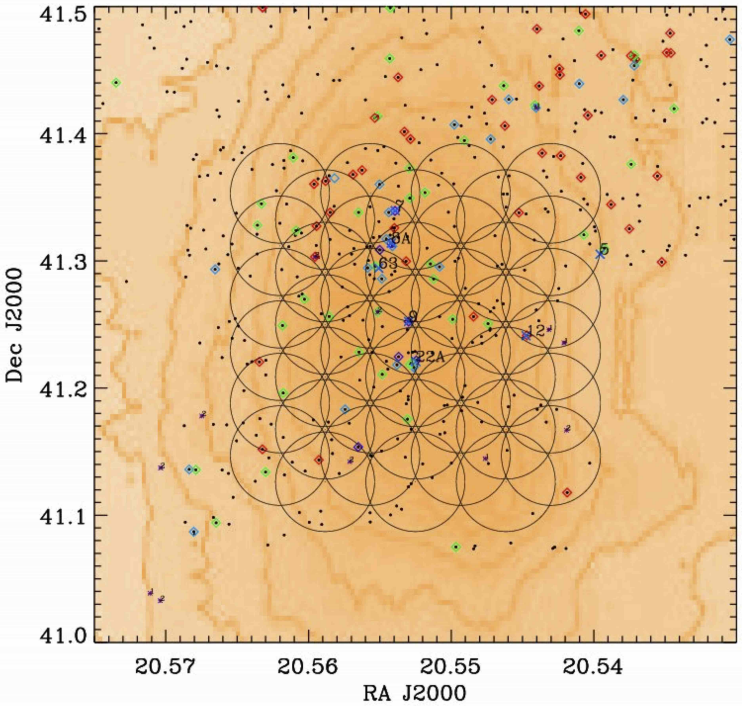
2. *Calibration* - standard AIPS routines will be used to calibrate the data, however the use of the e-MERLIN pipeline (Argo, 2014) written in ParselTongue will be used whenever possible in order to optimise this lengthy procedure. Key processes include the removal of any Radio Frequency Interference (RFI) that requires the use of various data visualisation methods, followed by the phase, amplitude and bandpass calibration.

3. *Imaging* - the wide-band coverage of e-MERLIN (especially at L-band) will necessitate the use of wide-field imaging routines for each individual pointing. Furthermore, these wide-field images will be mosaiced together, a challenging process for a radio interferometric array such as e-MERLIN with antennas that differ in size.

4. *Source Detection* - a novel source extraction algorithm, (see Section 3.1.1) will be implemented in order to detect and measure the integrated flux densities and positions for each source found within the COBRaS data. Once compiled, the COBRaS




(A) L-Band (21cm)




(B) C-Band (6cm)

FIGURE 1.14: The COBRaS positions for the (a) seven L-band (21cm) and (b) 42 C-band (6cm) pointings to be imaged with e-MERLIN according to the primary beam of the Lovell antenna. The background figure is the outline of the Cygnus OB2 association as given in Knödlseder (2000) with black dots representing counts from the 2MASS survey (Cutri et al., 2003). The blue symbols indicate some of the well known OB and WR stars within the stellar association.

catalogue can be cross-correlated with existing catalogues covering the Cygnus OB2 assocation.

5. *Data Archiving* - the final COBRaS source catalogue containing all measured quantities will be made readily available to the community within the existing Virtual Observatory (VO) format.

## 1.7 The aims and structure of this thesis

This thesis aims to firstly outline the theoretical framework across a number of astrophysical subjects including young massive clusters, radio interferometry, the mass-loss from massive stars and their stellar winds, colliding wind binaries and the incidence of non-thermal radiation. Moreover, in the reduction and scientific analysis of 21cm (L-band) observations obtained with the e-MERLIN interferometer (as part of the COBRaS Legacy survey), a number of key issues across these fields will be addressed. Furthermore, the technical work carried out in this process will not only aid the future work carried out within the COBRaS project, but also the wider community, appealing to the users of present and future radio telescopes such as e-MERLIN, MEEKat and the SKA.

Chapter 2 describes in detail the work undertaken in the reduction of the COBRaS L-band (21cm) data following a brief introduction to interferometric theory. This chapter addresses the current and increasingly important issue of Radio Frequency Inference (RFI) within the field of radio astronomy. Some of the techniques used to identify and remove the RFI are described, and some ideas for its future treatment are discussed. Additionally, the challenges faced in the calibration and imaging of the COBRaS L-band dataset are revealed through a detailed description of entire data reduction process.

Chapter 3 presents the work done in compiling the entire source population catalogue of these reduced COBRaS L-band radio maps. Along with the final catalogue, the chapter presents a new source detection algorithm SEAC (Source Extraction Algorithm for COBRaS), designed specifically to detect sources from the COBRaS datasets, measure their positions on the sky and calculate their integrated flux densities. The most recent radio

survey of the Cyg OB2 association at 21cm was compiled using the Westerbork Synthesis Radio Telescope (WSRT; Setia Gunawan et al., 2003), reaching a $5\sigma$ flux-density limit of $\sim$ 2 mJy. The COBRaS e-MERLIN L-band observations presented here significantly improve on that, in reaching a $5\sigma$ sensitivity limit of $\sim$ 100 - 125 $\mu$Jy in the centre of each pointing. The detected sources offer a direct comparison to those detected in previous surveys, providing further information of the source spectral index and perhaps the nature of previously unidentified sources. The final COBRaS L-band all source catalogue (CLASC) will serve as a template for the much anticipated COBRaS C-band observations, due to commence in early 2017.

The remaining chapters focus on the scientific analysis of the COBRaS L-band dataset, addressing the two main scientific goals behind the entire COBRaS project (see Section 1.6.1). Chapter 4 details the thermal free-free emission from single massive star winds, aiming to place constraints on the mass-loss and clumping within the winds of these hot stars. Drawing from the discussion made in Chapter 1 regarding the science behind OB star winds, their currently accepted model and the impacts that small and large scale structures have on their wind parameters and subsequent evolution, Chapter 4 describes the tools necessary to predict $\dot{M}$ from the thermal free-free emission from their winds. Using the results of the COBRaS L-band dataset, this chapter addresses the current discrepancy between mass-loss rates inferred through different diagnostics.

Chapter 5 focuses on the non-thermal emission from the association's massive star binaries. The longer wavelengths sampled at L-band in comparison to C-band will favour the detection of non-thermal emission due to its negative spectral index in the 'normal' convention (i.e. increasing flux with decreasing frequency/increasing wavelength). As such, any synchrotron emission from the colliding winds of massive stars will be more detectable at 21cm. This chapter explores the mechanisms driving the production of synchrotron emission from the colliding winds of two massive star binaries. In conjunction with the spectroscopically derived properties of Cyg OB2 massive stars binaries (Kobulnicky et al., 2014), the detections of such systems in the COBRaS L-band dataset will allow for an investigation into the dependance of the synchrotron emission on both the orbital parameters and those of the individual stars.

This thesis concludes with a summary of its content and its major scientific results in Chapter 6. The final discussion highlights the intended future work of the COBRaS project, whilst looking at both the technical aspects and further science that will be achieved in the advent of the C-band dataset. This thesis will serve not just as a useful template for the work to be carried out upon the COBRaS C-band dataset, but also as a valuable resource regarding the content of the Cygnus OB2 association at 21cm, showcasing the vast amount of stellar astrophysics that can be achieved through radio astronomy.

# Chapter 2

# e-MERLIN COBRaS L-band data treatment and reduction

The purpose of this chapter is two-fold. Firstly as a guide, outlining the basic theory behind radio interferometric observations and secondly as a reference, detailing the precise treatment and reduction of the e-MERLIN 21 cm (L-band) observations obtained as part of the COBRaS Legacy project. After an overview of interferometric theory, Section 2.2 will document the removal of Radio Frequency Interference (RFI) from the 21 cm data set, an essential task to enable accurate calibration of the data. Section 2.3 will give a thorough explanation of the steps taken to calibrate the L-band data set before Section 2.4 details the various imaging procedures used to create the final COBRaS 21 cm radio maps of the Cyg OB2 association.

## 2.1 An introduction to radio interferometric theory

A single parabolic dish or antenna can be used to measure the voltage of an astronomical signal at radio wavelengths. The parabolic nature of the dish is such that the antenna is sensitive to radio wavelengths from a particular direction. As a result, each antenna will have an associated primary beam pattern that describes the sensitivity of the antenna to a given direction. An antenna's sensitivity is at a maximum at the phase centre but drops off in directions further away from the phase centre. An example of a typical primary beam pattern for a parabolic antenna is shown in Figure 2.1. The pattern is comprised of

a main lobe centred around the direction to which the antenna is pointing and various side lobes, which show their sensitivity to unwanted radiation from undesired directions. The primary beam pattern must be corrected for when observing astronomical sources with such instruments, a point returned to in Section 2.1.2.

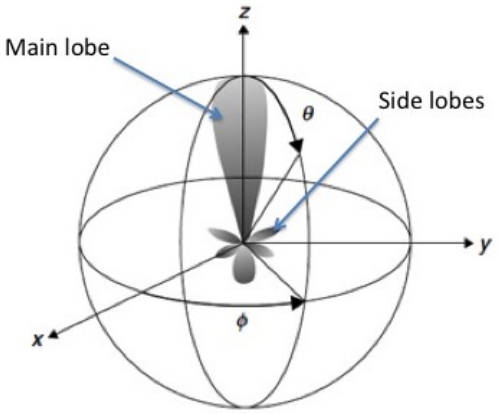



FIGURE 2.1: A cartoon depicting the response or primary beam pattern of a given antenna, pointings in the $z$ direction.

Interferometry is somewhat of an umbrella term that broadly describes the technique involving the superposition of waves (usually electromagnetic). The result of their superposition (or combination) will contain information regarding the initial state of the individual waves. Within radio astronomy, the interferometric technique used is more specifically known as 'Aperture Synthesis Imaging' and was founded by Antony Hewish and Martin Ryle of Cambridge University, for which they received a Noble prize in 1974 (along with the discovery of pulsars). By the use of two or more antennas that collectively observe an astronomical source at a given frequency, the combination of their signals offer a resolution equivalent to that of an antenna of diameter equal to the largest distance between its separate elements. For an antenna with an aperture of diameter D and wavelength $\lambda$, the resolution of the produced image is:

$$\theta" \sim \frac{2\lambda_{cm}}{D_{km}} \tag{2.1}$$

When observing for example, at a wavelength of 21 cm, to obtain a resolution of 1" we require an antenna aperture of $\sim$ 42 km. The production of a single dish antenna of this

size is simply not feasible. In order to significantly increase the resolution of images at such long (radio) wavelengths, one must synthesise a much larger aperture through summations of antenna pairs. This technique is known as 'Aperture Synthesis Imaging' and utilises the rotation of the Earth and the summation of antenna pairs to synthesise an aperture with a diameter equal to that of the largest distance between one of the antenna pairs. This revolutionised the field of radio and sub-mm astronomy, with the ability to produce images of sub-arcsecond resolution at such wavelengths.

### 2.1.1 The two element interferometer

Consider the simplistic case of the two element interferometer whereby two antennas of equal diameter are separated by a distance $b$ that is referred to as the baseline length (see Figure 2.2). Each antenna points towards an astronomical source assumed to be at an infinite distance from the interferometer. Under this assumption, the electro-magnetic wavefront collected at each antenna will be flat. The incoming radiation is also assumed to be quasi-monochromatic such that the wavefront is confined to one wavelength, the observing bandwidth of the antennas $\Delta\nu$ is also assumed to be very small. As such, the temporal dependance of the electric field of this incoming radiation can be written as:

$$E_\nu(t) = Acos(2\pi\nu t + \phi) \tag{2.2}$$

The amplitude $A$ and phase $\phi$ will remain unchanged to a time duration of the order $dt = 1/d\nu$ since the electric fields will be sinusoidal within this time frame. With reference to Figure 2.2, the geometry of the two element interferometer can be seen. The incoming wavefront from the direction of the source described by the unit vector $\vec{s}$, is collected at both elements separated by baseline $\vec{b}$. This signal at antennas 1 and 2 are measured as a voltage, $V_1$ and $V_2$, at their individual arrival time $t$.

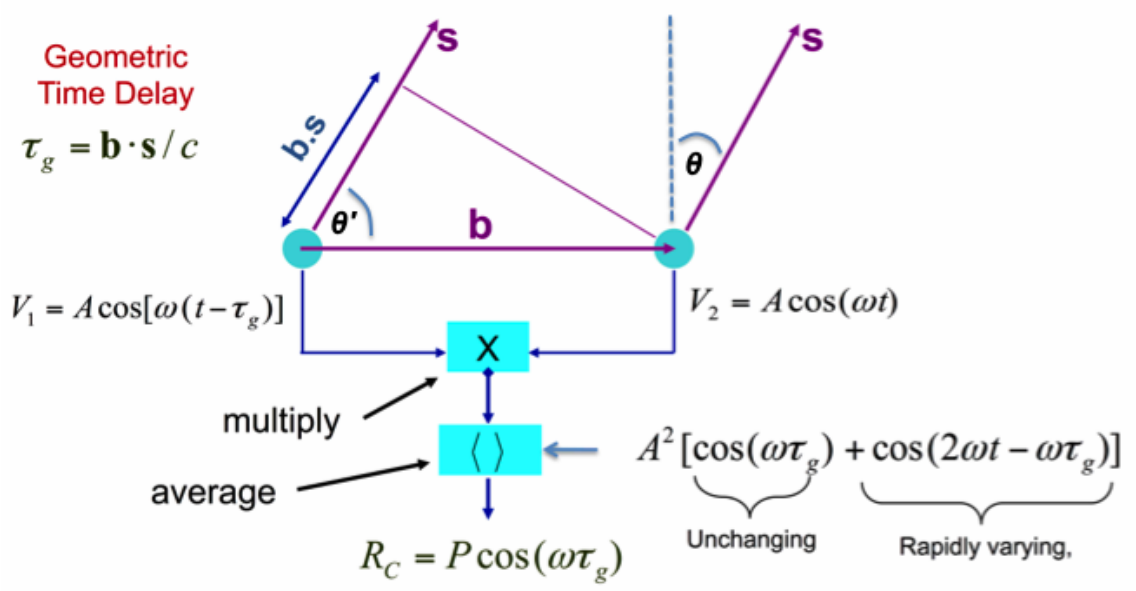



FIGURE 2.2: A schematic illustration of a pair of antennas separated by a baseline of length $b$ receiving a wavefront from the direction of the unit vector $\vec{s}$. The geometric time delay between the wavefront hitting the first antenna to the second $\tau_g = \frac{\vec{b}\cdot\vec{s}}{c}$. Antennas 1 and 2 record the signal voltage $V_1$ and $V_2$ respectively which are then multiplied together and averaged to produce the cosine correlator response, $R_c$.

We can define the geometrical time delay as the time between the arrival of the wavefront at antenna 1, and the arrival of the same wavefront at antenna 2 as:

$$\tau_g = \frac{\vec{b}\cos\theta'}{c} = \frac{\vec{b}\cdot\vec{s}}{c} \tag{2.3}$$

where $c$ is the speed of the light and $\theta' = (\pi/2 - \theta)$. With reference to Equation 2.2 the voltage measured at each antenna can be written as $V_1(t) = A\cos(2\pi\nu(t-\tau_g))$ and $V_2(t) = A\cos(2\pi\nu t)$, which are then multiplied together and averaged to obtain the cosine correlator response, $R_c$. In this quasi-monochromatic radiation example this equates to

be:

$$\begin{aligned} R_c &= \langle V_1(t) \times V_2(t) \rangle \\ &= \langle A cos(2\pi\nu(t-\tau_g)) \times A\cos(2\pi\nu t) \rangle \\ &= \langle \frac{1}{2}A^2[\cos(2\pi\nu(t-\tau_g)+2\pi\nu t) + \cos(2\pi\nu(t-\tau_g)-2\pi\nu t)] \rangle \\ &= \langle \frac{1}{2}A^2[\cos(4\pi\nu t - 2\pi\nu\tau_g) + cos(2\pi\nu t)] \rangle \\ &= I_\nu \cos(2\pi\nu\tau_g) \end{aligned} \tag{2.4}$$

where the rapidly varying term $\cos(4\pi\nu t - 2\pi\nu\tau_g) \rightarrow 0$ and the trigonomic formula $\cos(\theta)\cos(\phi) = 1/2(\cos(\theta+\phi)+\cos(\theta-\phi))$ has been used. Note that $R_c$ is a function of the intensity $I_\nu$ ($= \frac{1}{2}A^2$) and the geometric time delay, $\tau_g$ (and hence the orientation of the baseline). $R_c$ is not a function of the baseline location, the time of the observations nor the phase of the incoming signal.

We must account for the power received by the correlator over all parts of the sky by integrating $R_c$ over $4\pi$ steradians or solid angle $d\Omega$. In a similar fashion of multiplying the voltage signal from either antenna and averaging, this obtains the cosine correlator response:

$$R_c = \langle \int_\Omega V_1(t)d\Omega \times \int_\Omega V_2(t)d\Omega \rangle = \int_\Omega I_\nu(\tau_g) cos(2\pi\nu\tau_g) d\Omega \tag{2.5}$$

where $I_\nu(\tau_g)$ is now the source brightness on the sky. Currently this equation only takes into account the emission from the source with even symmetry due to its co-sinusoidal nature. In order to reveal the odd component, the correlator must also be sensitive to those fringes with a phase difference of $\pi/2$. This is accomplished by inserting a 90° phase shift into the signal paths of one of the antennas to obtain a sinusoidal fringe pattern, i.e. $V_2 = A_2\cos(2\pi\nu t - (\pi/2)) = A_2\sin(2\pi\nu t)$. Running this through the same process of multiplying together the two voltages and averaging obtains the sine correlator response $R_s = \int_\Omega I_\nu(\tau_g)\sin(2\pi\nu\tau_g)d\Omega$. From the two independent correlator outputs $R_c$ and $R_s$,

one can define the complex visibility function:

$$\begin{aligned} V(\tau_g) &= R_c - iR_s = Ae^{-\phi} \\ &= \int_\Omega I_\nu(\tau_g)\left(\cos(2\pi\nu\tau_g) - i\sin(2\pi\nu\tau_g)\right) d\Omega \\ &= \int_\Omega I_\nu(\tau_g) e^{-i2\pi\nu\tau_g} d\Omega \end{aligned} \tag{2.6}$$

where the amplitude $A = \sqrt{R_c^2 + R_s^2}$, the phase $\phi = tan^{-1}\frac{R_s}{R_c}$ and Euler's formula $e^{-i\phi} = \cos(\phi) - i\sin(\phi)$ has been used, with $\phi = 2\pi\nu\tau_g$. This is a 2-D Fourier transform that allows us to establish the real function $I_\nu(\tau_g)$ from the complex, hermition function $V(b)$. The term visibility, first used by Michelson (1890), is a complex quantity that describes the relative amplitude of the observed fringes, the magnitude of which has units of spectral power flux density ($\mathrm{Wm^{-2}Hz^{-1}}$; Thompson 1999). Equation 2.6 is otherwise known as the measurement equation and is shown to be a function of the time delay and hence the baseline orientation. We now introduce the $u, v, w$ axes, a coordinate system describing the baseline vector $\vec{b}$ and the direction cosines $l, m, n$ that define the projection of the unit vector $\vec{s}$ on the sky as follows:

$$\nu\tau_g = \frac{\nu\vec{b}\cdot\vec{s}}{c} = ul + vm + wn \tag{2.7}$$

Similarly, it follows that the solid angle on the surface of the sky, $d\Omega = dldm/n$. Substituting these expressions into Equation 2.6 obtains the visibility function in terms of these coordinate systems, showing the relation between the measured visibilities as a function of baseline orientation in terms of the $u, v, w$ coordinates and the observed source brightness as a function of the sky directional cosine coordinates $l, m, n$.

$$V(u,v,w) = \int_l \int_m \int_n I_\nu(l,m,n) e^{-i2\pi(ul+vm+wn)} \frac{dldm}{n} \tag{2.8}$$

where $n = \sqrt{1 - l^2 - m^2}$. This is a 3-dimensional Fourier transform equation in the simplified case of the two element interferometer. Each element measures a voltage from an

astronomical signal at frequency $\nu$, which are correlated together. The measured complex visibility function is obtained combining the cosine and sine interferometer response as shown in Equation 2.6. In defining a new coordinate system as defined by Equation 2.7, this function can be re-expressed as shown in Equation 2.8 and the source brightness $I_\nu(l, m, n)$ can be obtained by performing an inverse Fourier transform upon the measured visibility function $V(u, v, w)$.

### 2.1.2 A realistic radio interferometer

Taking a more realistic approach, we must relax the assumption of quasi-monochromatic radiation. In reality, interferometers observe over a finite bandwidth or frequency range $\Delta\nu$. To compensate we must integrate the interferometer response over this bandwidth $\Delta\nu$ centred at frequency $\nu_0$ as follows:

$$\begin{aligned} R_c &= I_\nu \int_{\nu_0-\nu/2}^{\nu_0+\nu/2} \cos(2\pi\nu\tau_g) d\nu \\ &= I_\nu \,\mathrm{sinc}(\Delta\nu\tau_g) \cos(2\pi\nu\tau_g) d\nu \end{aligned} \tag{2.9}$$

The introduction of the $\mathrm{sinc}(x)$ function, often referred to as the 'bandwidth pattern', attenuates the source emission (Thompson, 1999). This attenuation will be significant if the source offset from the phase centre is comparable to the interferometer resolution divided by the fractional bandwidth. Thus when observing over a large bandwidth it is best to observe over multiple, narrow channels in order to avoid such 'bandwidth attenuation'.

A further consequence of observing over a finite bandwidth is that the sensitivity of the interferometer is not uniform over the sky, i.e. sources not on the plane orthogonal to the interferometer baseline will not be observed with full coherence. Therefore, it is necessary to add a time delay in order to shift the maximum of the 'sinc' pattern (or fringe attenuation pattern) to the centre of the source. For reference Figure 2.3 shows a schematic of the additional time delay $\tau_0$ in the simplest case of the two element interferometer. In running

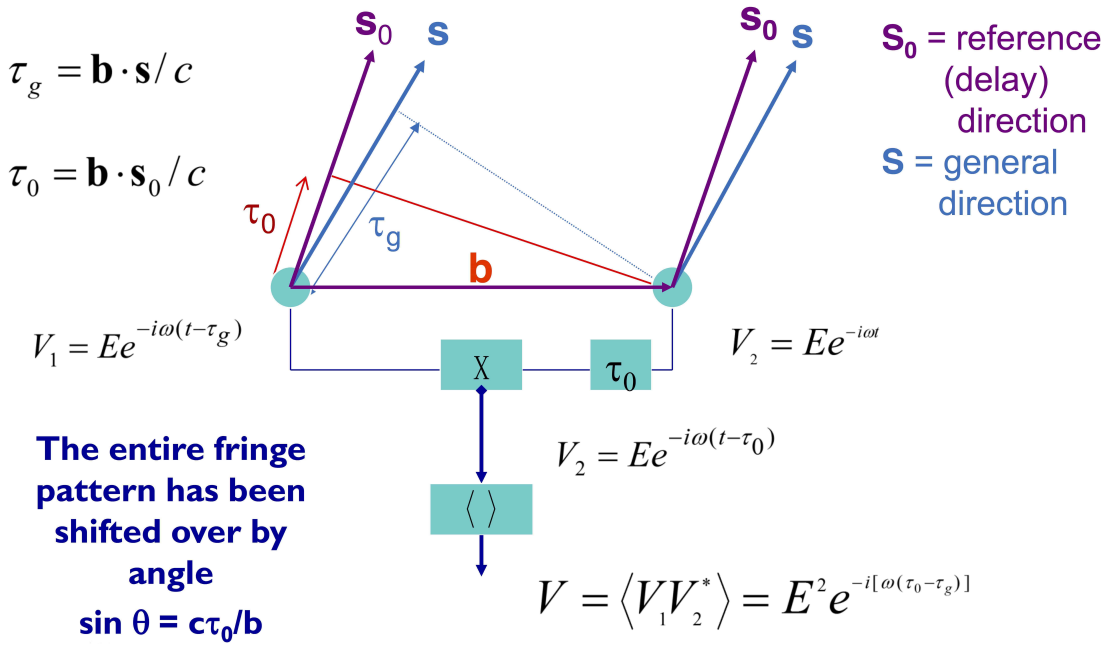



FIGURE 2.3: A schematic illustration of a pair of antennas separated by a baseline length $b$ receiving a wavefront from the general direction described by the unit vector $\vec{s}$ as shown in Figure 2.2. This schematic also shows the additional time delay, $\tau_0$ added to the received signal at each antenna to shift the fringe attenuation pattern to the reference (delay) direction described by the unit vector $\vec{s}_0$.

through the same steps in deriving the complex visibility function as before but with this additional time delay gives:

$$\begin{aligned} V(\tau_g) &= R_c - iR_s \\ &= \int_\Omega I_\nu(\tau_g) e^{-i2\pi\nu(\tau_0-\tau_g)} d\Omega \end{aligned} \tag{2.10}$$

It is necessary to continuously track the source by changing this time delay $\tau_0$ such that the fringe pattern moves along with the source as the Earth rotates. This must be done to an accuracy of $\delta\tau << 1/\Delta\nu$ in order to minimise the effect of bandwidth attenuation. In expressing the complex visibility function defined in Equation 2.10 as a function of the $u, v, w$ and $l, m, n$ coordinate systems where the $w$ axis points towards the reference centre $\vec{s}_0$, we find that:

$$\nu\tau_0 = \frac{\nu\vec{b}\cdot\vec{s}_0}{c} = w \tag{2.11}$$

Inserting this expression, in combination with Equation 2.7, into Equation 2.10 means that the 3-D Fourier transform equation becomes:

$$V(u,v,w) = \int_l \int_m \int_n I_\nu(l,m,n) e^{-i2\pi[ul+vm+w(\sqrt{1-l^2-m^2}-1)]} \frac{dldm}{\sqrt{1-l^2-m^2}} \tag{2.12}$$

In order to simplify the inverse Fourier transform of this equation, whereby $I_\nu(l,m)$, the source brightness on the sky is obtained from the measured visibility, it is often reduced into the form of a 2-D Fourier transform. This is done under the condition that $w = 0$ when the baselines are considered to be coplanar, or such that $l$ and $m$ are sufficiently small such that $w(\sqrt{1-l^2-m^2}-1) \approx 0$. Thus, Equation 2.12 reduces to:

$$V(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I_\nu(l,m) e^{-i2\pi(ul+vm)} dldm \tag{2.13}$$

The interferometer observes visibilities $V_\nu(u,v)$ from which one can inverse Fourier transform Equation 2.13 to obtain the observed sky brightness of the source in question, $I_\nu(l,m)$. However, note that the technique used in radio interferometry is known as 'aperture synthesis imaging' because the interferometer synthesises an aperture of diameter equal to that of its largest baseline. Thus, when an interferometer observes a source it is only sampling part of the sky, since the visibility function is measured at particular places in the $u,v$ plane. This sampling is described by the sampling function $S(u,v)$ that equates to zero at those $u,v$ coordinates where no data is taken. The Fourier transform of the sampled visibility data is known as the 'dirty image' and is written mathematically as:

$$\begin{aligned} I_\nu^D(l,m) &= \int \int V_\nu(u,v) S(u,v) e^{2i\pi(ul+vm)} dudv \\ &= I_\nu(l,m) \bigotimes B(l,m) \end{aligned} \tag{2.14}$$

where the bottom half of above equation states that the true image of the sky $I_\nu(l,m)$, when convolved with the synthesised beam $B(l,m)$ (note that the symbol $\bigotimes$ denotes convolution) gives the dirty image $I_\nu^D(l,m)$. The synthesised beam, $B(l,m)$ is often referred to as the Point Spread Function (PSF) or indeed the dirty beam. It follows that whilst the inverse

Fourier transform of Equation 2.13 yields the true sky image $I_\nu(l,m)$, the synthesised beam $B(l,m)$ is obtained via the inverse Fourier transform of the sampling function, i.e.

$$B(l,m) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} S(u,v)e^{-\pi(ul+vm)}dudv \tag{2.15}$$

The rotation of the Earth is what is used to improve the $u,v$ coverage of the sampling function. As the Earth rotates around its axis, the orientation of each baseline associated with the interferometer changes with respect to the source, filling out the $u,v$ plane. This results in lower side-lobes (see Figure 2.1) within the sampling function and produces a better image on the sky. The most common method used for the imaging and deconvolution process is the CLEAN algorithm (see Section 2.4.1) although others such as the Maximum Entropy Method (MEM) also exist.

Another assumption in need of relaxing is that our antennas are equally sensitive to all parts of the sky. In reality, antennas are point probes that simply measure the voltage of the incoming electromagnetic waves at a specific point on the sky. Since, they are dishes, they are directional elements and in order to produce accurate images of the sky, it is necessary to include a factor, $A_\nu(l,m)$ that describes the normalised antennas response pattern as a function of the directional cosines $l,m$ (i.e. the primary beam pattern, see Figure 2.1). If all the antennas of the interferometer are identical and their reception patterns are constant in time, the antennas response $A_\nu(l,m)$ can be inserted into the 2-D Fourier transform equation (i.e. Equation 2.13):

$$V_\nu(u,v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} A_\nu(l,m)I_\nu(l,m)e^{-2i\pi(ul+vm)}dldm \tag{2.16}$$

Thus, the final sky image is obtained by dividing the true image of the sky $I_\nu(l,m)$, by the antenna response pattern $A_\nu(l,m)$. The above outlines the theoretical framework behind synthesis imaging. The following sections will describe the more specific steps taken in the treatment, calibration and imaging of the e-MERLIN COBRaS L-band data set.

## 2.2 Radio frequency interference

Radio Frequency Interference (RFI) describes the effect of unwanted energy due to some form of emission, radiation or induction upon the reception of a radio communication system. In regards to radio astronomy, this inference manifests itself in signals generated from various types of man-made electronic equipment. RFI is one of the main factors limiting the performance of radio telescopes by constricting the frequency space available, corrupting astronomical signals and introducing errors into the calibration and subsequent radio images. It has become an ever increasing problem within the community as a result of technical advancements that are constantly increasing the sensitivity of radio telescopes. Large increases in the bandwidth used by modern interferometers coupled with other technological advances (e.g. mobile phones and wifi) utilising the same frequency space has led to the development of new methods in order to mitigate such interference.

RFI mitigation techniques can be implemented at two different stages of the observational process, namely pre- and post-correlation. Whilst methods applied at the pre-correlation stage can offer observations at their highest time resolution (i.e. without the need to remove corrupted chunks of the data), hardware constraints mean they are often limited to handling small time ranges or the data from a single dish (Offringa et al., 2010). Several methods at the pre-correlation stage do exist however, such as a threshold method utilising $\chi^2$ statistics from Weber et al. (1997), or the cumulative sum method (CUSUM) developed by Baan, Fridman, and Millenaar (2004). Post-correlation methods have traditionally relied on manual flagging techniques to remove corrupted data. More recently, automated post-correlation mitigation methods include: the use of an independent RFI reference signal in order to subtract the RFI (Briggs, Bell, and Kesteven, 2000), fringe fitting (Athreya, 2009), and a sum threshold technique (Offringa et al., 2010; Peck and Fenech, 2013). These various mitigation techniques are only applicable in some circumstances, as many of them have been developed specifically for use with data from a particular telescope or to deal with a particular type of RFI.

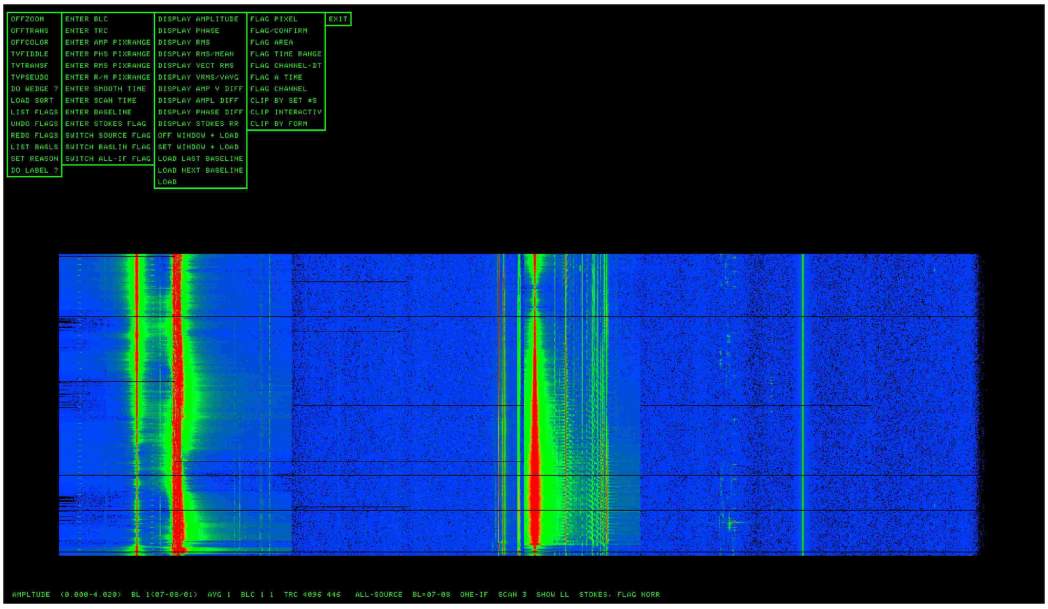
RFI manifests itself in many different forms (Lemmon, 1997; Fridman and Baan, 2001) and can vary significantly in intensity across time and frequency space. The sources of RFI span

a huge range from mobile phones, CCTV cameras, aeroplanes, microwaves, satellites and even through the internal electronics of the instrument. The combination and variability of such diverse sources of radio transmission give rise to the unpredictable nature of radio frequency interference. Moreover, each radio telescope has a different location, baseline orientation, frequency band, sensitivity limit and RFI environment. All of which must be considered when implementing RFI mitigation techniques to suit a particular instrument.

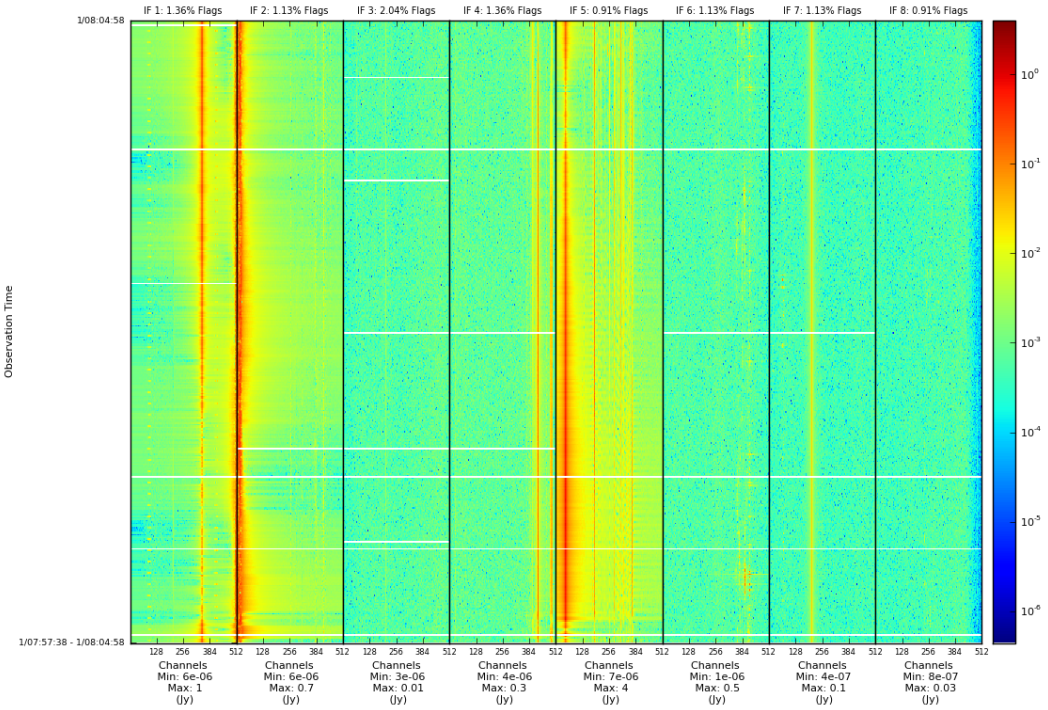
In the case of the e-MERLIN array based at Jodrell bank, there currently exists no hardware in order to tackle pre-correlation RFI mitigation. Therefore, its treatment and removal must be applied at the post-correlation stage. In practice, this is accomplished using the combination of an automated, SumThreshold method mitigation procedure and by visually inspecting the data in order to locate and manually flag the RFI using a variety of different AIPS tasks. The following section (2.2.1) will describe a new data visualisation programme, SPectral Plot (SPPlot), used to visually inspect radio visibility data with the primary aim of easily identifying RFI. Section 2.2.2 will then briefly describe the Scripted E-MERLIN RFI-mitigation PypelinE for iNTerferometry (SERPent; Peck and Fenech 2013) used in the treatment of e-MERLIN data. Section 2.2.3 will then describe how both SERPent and SPPlot were implemented in the removal of RFI of the COBRaS L-band data set, prior to its calibration.

### 2.2.1 SPPlot

Radio interferometric data can be viewed or inspected in a huge number of different ways. The Astronomical Image Processing System (AIPS) used for the reduction and imaging of the COBRaS L-band data sets contains numerous ‘tasks’ that allow the user to plot or view the data across various parameter spaces. In terms of successfully identifying RFI at the post-correlation stage, the data is best inspected in 3-dimensions, with frequency, time and amplitude as the respective $(x, y, z)$ axes. The inbuilt AIPS task ‘SPFLG’ allows the user to plot their radio visibilities in this way and highlight sections of the data that can be subsequently flagged and written into an AIPS flag (FG) table. The task gives the user little freedom over how the data is plotted onto the screen and due to the large bandwidth




(A) SPFLG Interface




(B) SPPlot Output

FIGURE 2.4: To highlight the difference between (a) the graphical interface from the AIPS task SPFLG and (b) an output plot from the data visualisation programme SPPlot. Both plots show the same chunk of data, although SPFLG only displays every three channels in the frequency direction.

used in modern radio telescopes, AIPS will often average the data in order to fit everything onto the screen. This gives an inaccurate representation of the radio visibilities and leads to less reliable identifications of RFI. Furthermore, SPFLG has no option to save the 3-D plots and only allows the user to view the images in an interactive format, displaying only one polarisation of a single baseline at any one time.

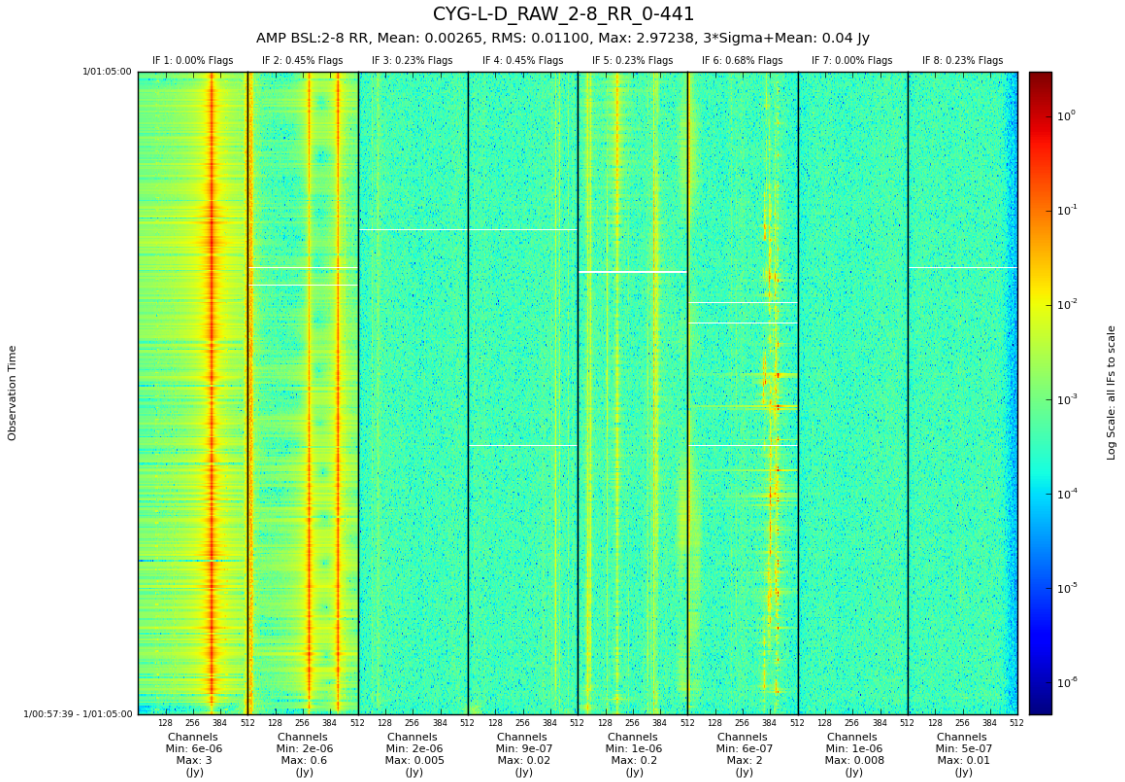



FIGURE 2.5: An output from SPPlot showcasing the more detailed format option enabling for a better inspection of RFI. The SPPlot shows a completely raw data chunk of the RR polarisation of a single scan from baseline 2-8 (MK2 to Darnhall) of the COBRaS Legacy L-band D pointing taken on 2014 Apr. 26th. Strong broadband RFI can be seen across IFs 1 and 2 while some narrow band components can be seen in IF5.

SPectrum Plot (SPPlot), a data visualisation programme, was written with the aim of re-creating the 3-dimensional plot of SPFLG (see Appendix A for further details) yet giving the user far more freedom over the parameter space. A direct comparison between the outputs from SPPlot and SPFLG can be seen in Figure 2.4 where the same section of data has been used in both circumstances. SPPlot is written in the PYTHON/PARSELTONGUE environment for easy interaction with AIPS. It was originally based upon the PLOTLIB routines written by Hans-Rainer Kloeckner for the use on data from the LOw Frequency

ARray (LOFAR). The programme consists of a main script and an input file, allowing the user to easily view and manipulate the plotting parameters. The user has a large variety of options including which baselines, sources, polarisations, intermediate frequency (IF; AIPS speak for spectral window) and time range of the data to plot, as well as more aesthetic (i.e. plotting) specific parameters such as the colour scheme, image resolution and the z-axis colour scale (e.g. linear, logarithmic, etc). Furthermore, SPPlot will plot the phase or amplitude of the radio visibilities and load in, and apply a particular flag table associated with the data before plotting.

The programme itself is split into three separate sections. The first of which loops through and counts radio visibilities, obtaining the necessary information based upon the user defined parameters of the input file. This information is required for the programmes in-built functions and is stored in PYTHON dictionaries. The second section cycles through the desired radio visibilities, calculating either the amplitude or phase from the real and imaginary part of each visibility and stores the information in NUMPY (the fundamental package for scientific computing within PYTHON) arrays. In order to save vital computational time and memory, the amplitude or phase, calculated from each visibility, is placed into a NUMPY array corresponding to the baseline, IF, polarisation and time range of the visibility, which is then saved to the hard disk by the use of PYTABLES. If plotting a large section of data, the NUMPY arrays containing the visibility information can total up to hundreds of GigaBytes (GB) which is larger than the memory available in most machines. This process of saving the arrays to the hard disk is computationally in-expensive and absolutely necessary in the advent of large data sets. The third section of the programme then loads back in each NUMPY array and plots each page using PYTHONS plotting library MATPLOTLIB. Again, in order to be computationally efficient, the programme will plot one page at a time, meaning only a small amount of the memory is being used at any one time. Furthermore, this process is run in parallel, utilising the number of Computer Processing Units (CPUs) available which significantly speeds up the entire process.

The programme can output the created plots in two different formats. An example of the first can be seen in Figure 2.5, whereby each page is dominated by an entire plot of a single baseline, polarisation and time range. This format was designed to show the finer details

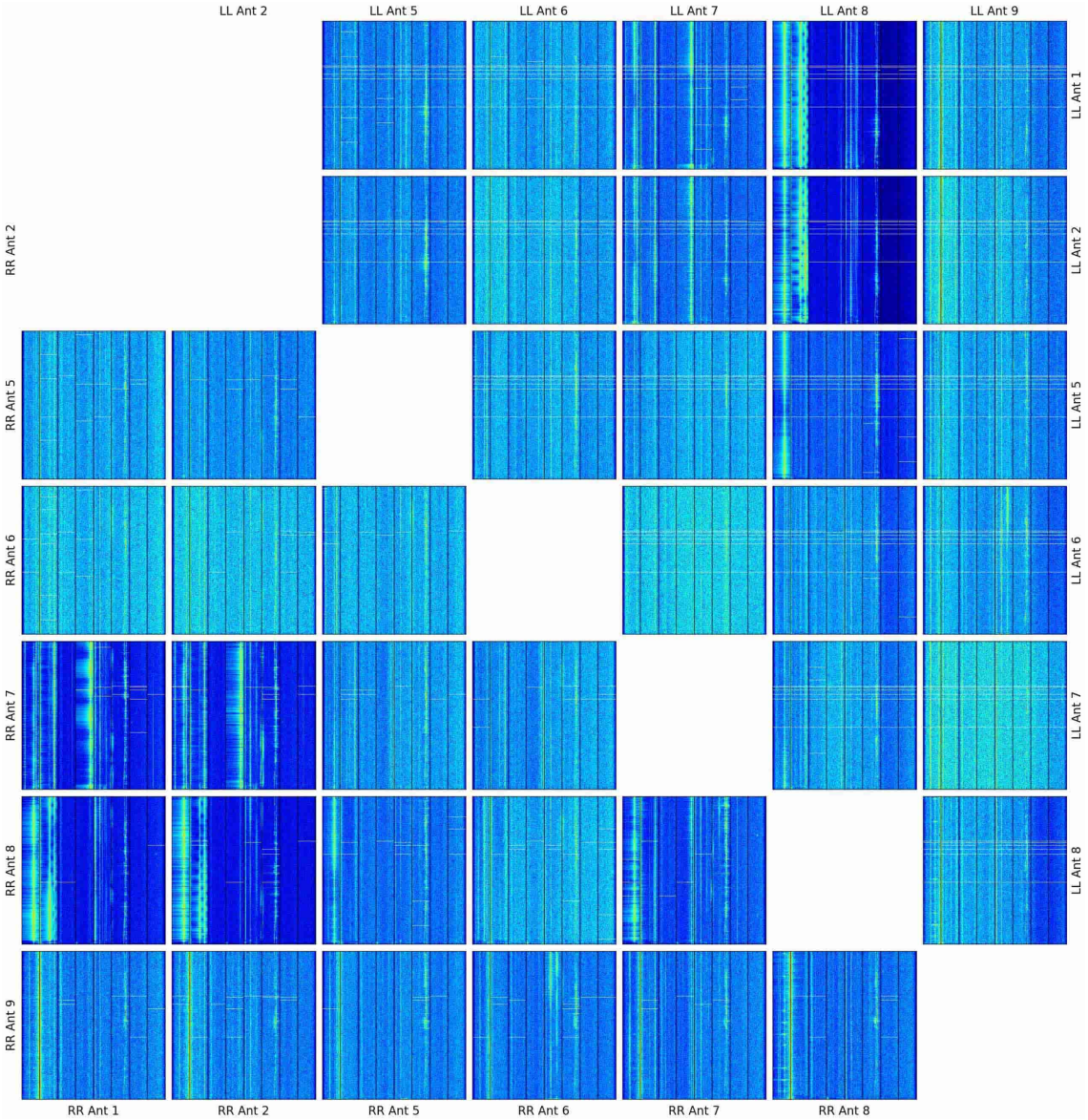



FIGURE 2.6: An output from SPPlot in its combined format. The plot shows the same data chunk as in Figure 2.5 however it shows all 20 available baselines across both the RR and LL polarisations. These plots are powerful diagnostic tools used to inspect the condition of radio visibilities throughout its acquisition, RFI treatment, calibration and pre-imaging.

such that RFI could be easily identified according to what time, channel and in which IF it appears in as well as its intensity relative to the astronomical signal. The second format (Figure 2.6) plots each baseline onto one single page designed to be a quick indicator as to the state of the radio visibilities[1]. SPPlot is therefore a useful tool throughout all stages of data reduction, from data acquisition and initial inspection, to the treatment and removal of RFI and throughout the entire calibration of the data. As such, the programme has now been incorporated into the e-MERLIN pipeline (Argo, 2014). SPPlot is made available to the public via GitHub and can be downloaded for free at the following web address - https://github.com/jackmorford/SPPlot.

### 2.2.2 SERPent

The scripted e-MERLIN RFI-mitigation pipeline for interferometry (SERPent) was written by Peck and Fenech (2013) at University College London (UCL). It was designed to offer an automated process in the mitigation of RFI within e-MERLIN data at the post-correlation stage, capable of efficiently handling data sizes on the order of TeraBytes (TBs). The programme utilises a SumThreshold methodology. Originally demonstrated by Offringa et al. (2010), it was found to be the most effective thresholding method currently available in tackling RFI at the post-correlation stage.

The SumThreshold method requires that the RFI visibility amplitudes are larger than those attributed to the astronomical source at the position in which they appear in time and frequency space. The RFI amplitudes must fulfil a certain threshold condition in which they are larger than some statistical level dictated by the relevant visibility subset in which they reside (i.e. the background or true source amplitudes). Since RFI can be independent of baseline, polarisation and IF, the size of this visibility subset from which the relevant statistics are calculated is very important. The two principle quantities that dictate the run of the programme are thus the threshold level $\chi$ and the subset size $N$.

[1]This format was originally conceived by Dr. Javier Moldon, currently based at the University of Manchester.

The threshold level $\chi(N)$ for some subset size $N$ is determined using the median and mean absolute deviation (MAD) of the sample and an 'aggressive' parameter $\beta$:

$$\chi(N) = median_i(x_i) + \beta MAD \tag{2.17}$$

where $x_i$ denotes the original sample. The subset size ($N$) describes the number of elements in a single dimension of the time and frequency array, of the visibility data that is to be tested. Within each subset, the visibility amplitudes are averaged together and if this average exceeds the calculated threshold level $\chi(N)$ of that subset, the elements are considered to be RFI and are flagged. SERPent first works through each permutation of the subset size $N$ of all elements in the time direction. The next threshold level is obtained considering those of the flagged elements via the equation:

$$\chi(N+1) = \frac{\chi(N)}{\rho^{(\log_2 N)}} \tag{2.18}$$

where typically, $\rho = 1.5$ (Offringa et al., 2010). The programme then moves on to the next subset size in a specified series, typically taken to be $S_N = [1, 2, 4, 8, 16, 32]$ to ensure a maximum efficiency between computational and flagging performance (Offringa et al., 2010). Having run through the time direction for each subset, SERPent repeats the entire process in the frequency direction, the completion of which constitutes one full run of the SumThreshold sequence. SERPent first completes an initial sequence with $N = 1$ in order to remove high amplitude RFI, then completes a further two runs to help eliminate remaining low-level RFI. A further parameter $\sigma_k$, known as the 'kickout' clause is included to prevent the algorithm from over-flagging the data set. If any threshold level $\chi(N) \leq mean(x_i) + \sigma_k \times MAD$, then the flagging run for that particular time or frequency direction is halted. For further details on the SumThreshold method, please refer to Offringa et al. (2010).

In the application of the SumThreshold methodology to e-MERLIN data, SERPent had to account for two problems, both a result of the fundamental characteristics of the array.

Firstly the largest antenna within the e-MERLIN array is the Lovell telescope with a diameter of 76m (Table 2.4). Relative to the other antennas in the array, it has a very slow slew speed which presents a problem unique to the e-MERLIN interferometer. Throughout an observing run, scans are alternated between the target and a bright phase reference source that are separated by no more than a few degrees (see Section 2.3.3 for further details). The Lovell antenna, with its slow slew speed only participates in observing the phase reference source every alternate scan of the phase referencing cycle. Therefore in every other phase reference scan, the Lovell remains on the target recording visibilities that are unusable and incoherent with the rest of the array. SERPent must identify and flag these Lovell stationary scans, which will be of a significantly lower amplitude than those spent on the phase reference source, before commencing the SumThreshold sequence. Secondly, zero-level (or off-source) visibilities are a characteristic of e-MERLIN data. Normally found at the beginning of each scan they can be a result of telescope slew errors, a system failure, or the recording of data before the telescope was physically 'on-source'. These are further visibilities that can skew the statistics of the SumThreshold algorithm and therefore must be flagged before the algorithm begins. For further details on how SERPent specifically deals with both the Lovell stationary scans and zero-level dropouts, please see Peck and Fenech (2013) and Peck (2014).

SERPent is continuously updated to overcome new problems encountered with different e-MERLIN data sets and correcting bug fixes. Major updates were released throughout 2015 in order to optimise the efficiency of the code and increase its over-all functionality. Like SPPlot, SERPent is a part of the official e-MERLIN pipeline (Argo, 2014) and is made publicly available from the following GitHub web address: https://github.com/daniellefenech/SERPent.

### 2.2.3 RFI removal process for COBRaS L-band data set

The COBRaS Legacy L-band observations were primarily performed over the course of 3 days between the 25th and 27th of April 2014 and constitute $\sim$ 75% of the entire data set. Additional observations were completed on the 11th of April 2014 to make up the remaining $\sim$ 25% of the awarded 42-hour allocation. Note from here on, these epochs will

be abbreviated to the Apr. 26th and Apr. 11th. Table 2.1 highlights some of the main characteristics of the entire COBRaS L-band Legacy data set. The observations cover 512 MHz of the electromagnetic spectrum, ranging from $\sim$ 1.25 - 1.77 GHz. This covers an area of the spectrum that is particularly prone to both narrow and broadband RFI. Some of the sources of RFI are known, such as that from CCTV cameras at a frequency of $\sim$ 1.376 GHz. However, the majority of the RFI picked up from e-MERLIN observations is almost entirely unpredictable and found to vary significantly in intensity over time and frequency space.

TABLE 2.1: Basic information regarding the COBRaS L-band Legacy data set. The total integration time and data size includes all of the seven target fields and the observed calibration sources over both observation epochs (Apr. 11th and Apr. 26th). Note that the given integration time and data size describes that of the raw data set, i.e. before any loss of data due to RFI, noisy data or 'off-source' visibilities.

| | |
|---|---|
| Total Integration Time | $\sim$ 48 hrs |
| Data size | $\sim$ 940 GB |
| Number of Pointings | 7 |
| Pointing FoV$^{\dagger}$ | $\sim 10.95'$ |
| Central Frequency | 1.5103374 GHz |
| Frequency Range | 1.2543374 - 1.7663374 GHz |
| Number of Baselines | 20 |
| Number of IFs | 8 |
| Number of Channels (per IF) | 512 |
| Channel Increment | 125 kHz |
| Number of Polarisations (Stokes) | 4 (RR, LL, RL, LR) |
| Predicted Resolution | 183 mas |
| Predicted Sensitivity | $\sim 10$ $\mu$Jy |

$^{\dagger}$As measured from the central frequency (listed) of the band

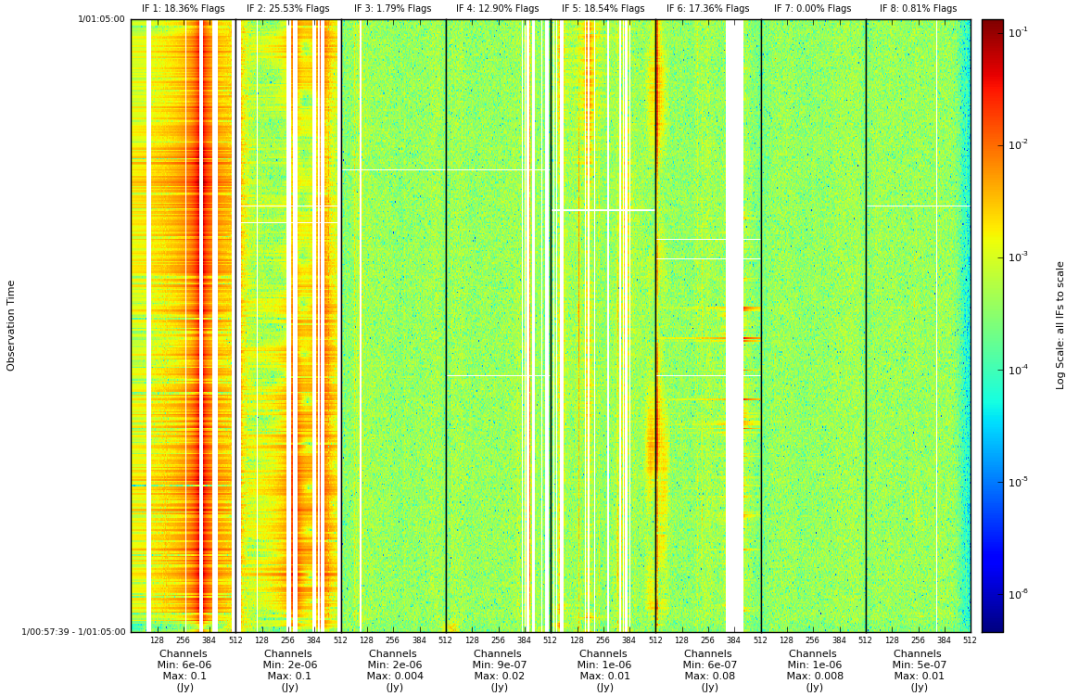
As part of the e-MERLIN pipeline (Argo, 2014), there exists a pre-written 'flagmask' which can be run as a standalone script. The script utilises the PARSELTONGUE environment to interact with the data within AIPS in order to create an AIPS FG table. Within the FG table are flags that correspond to the known narrow band RFI which occupies the L-band region of the radio continuum. This was the first flagging procedure applied to the COBRaS L-band data set and its application can be seen in an output from SPPlot in Figure 2.7a. For comparison, this is the same section of data shown in Figure 2.5 which is plotted prior to the flagmask application. These known narrow band RFI components equate to $\sim$ 12% of the data across the time and frequency space. The plot in Figure 2.7a

further reveals the vast extent of RFI (most notably in IF1) present in some sections of the e-MERLIN COBRaS 21cm observations. IF1 clearly shows how strong RFI can be at L-band, with the interference spilling across nearly all of the 512 channels of the IF.

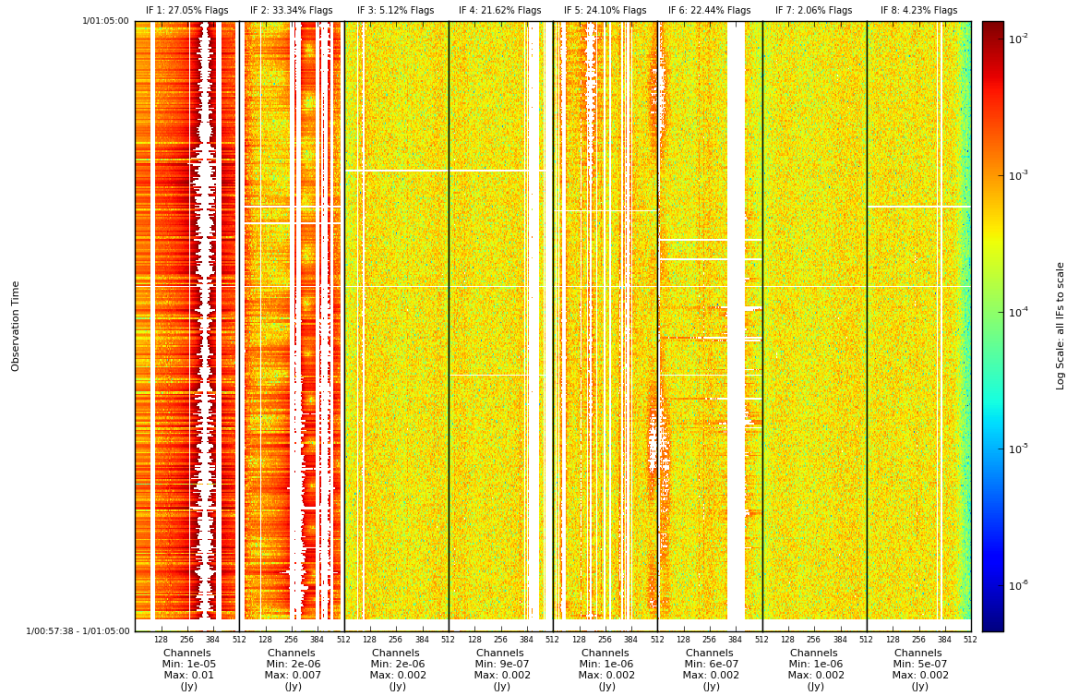
RFI such as that shown across IFs 1 and 2 of Figure 2.7a can be up to ∼3-4 orders of magnitude larger than the astronomical source. Furthermore, the RFI covers almost the entire channel width of IFs 1 and 2 rendering those visibilities corrupt and unusable, whereas other IFs, such as 7 contain almost no RFI. This disparity between levels of RFI between the different IFs proved to be an issue for the automated flagger SERPent as can be seen by Figure 2.7b. The flagging parameters for SERPent can be chosen as a function of source and baseline but cannot be varied from IF to IF. Thus aggressive parameters resulted in SERPent 'over-flagging' the data, whereby visibilities that belonged to the astronomical signal were flagged, whilst less aggressive parameters only partially mitigated the RFI (as seen in Figure 2.7b).

To overcome this problem, large sections of the data found to contain RFI, such as IFs 1 and 2 in Figure 2.7a, had to be manually flagged. A time consuming task upon a data set of nearly a terabyte in size. The data was first plotted within SPPlot in time order such that the phase-calibration scans were interspersed between the target field scans. In this way, the data could be inspected as a whole and flags could be written for large sections of the data where the RFI was almost constant in the time direction. The flags were written into a text file in such a format that it can be read into AIPS using the task UVFLG. This task converts the flags into an AIPS FG table which it attaches onto the data set. As such fewer flags were to be manually written into an AIPS FG table.

Once this manual flagging of the data set was completed for all polarisations, across all baselines and for all sources, the data was in a state where it could be run through SERPent (see Section 2.2.2). The flagging parameters for each source were chosen as a function of baseline and are shown in Table 2.2. Note that the parameter $\rho$ in equation 2.18, which defines the threshold of the next subset size, was set to 1.5 for all runs of SERPent. The flagging parameters were derived from a parameterisation study utilising several sections COBRaS L-band data that gave a fair representation of the entire data set. Following a




(A) e-MERLIN Flagmask Applied.

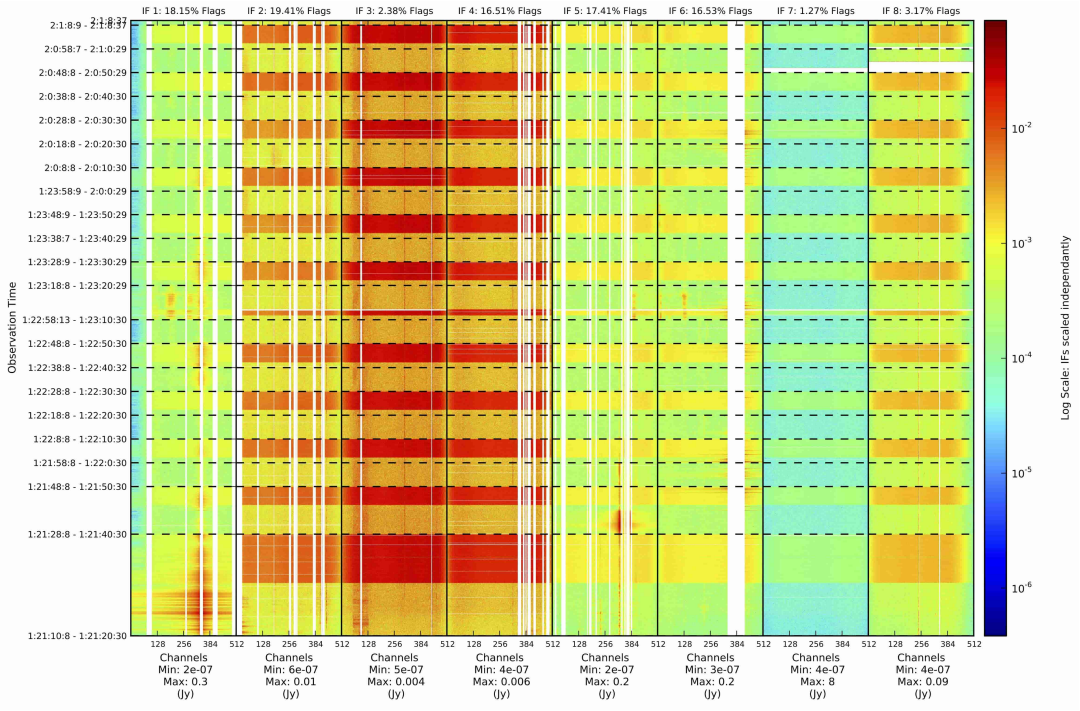



(B) Applied FG table from a typical SERPent run.

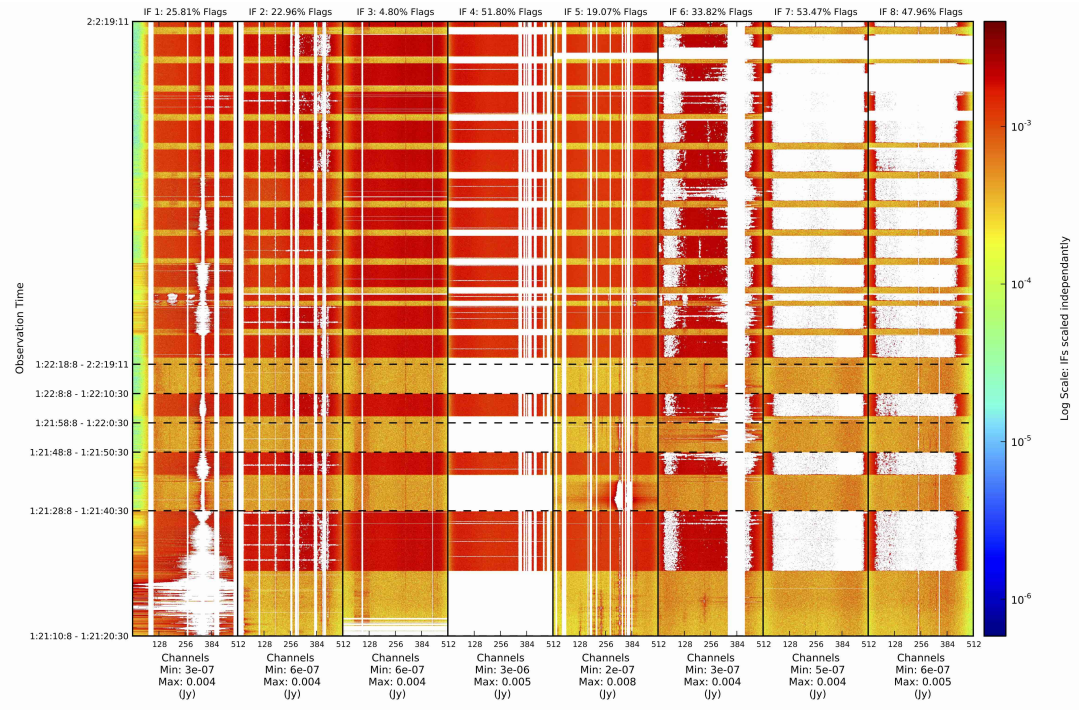
FIGURE 2.7: To highlight the application of (A) the e-MERLIN L-band flagmask and (B) a typical SERPent run on an example section of the raw COBRaS Legacy data taken on Apr. 26th. The visibilities in both cases belong to the same chunk of data as shown in Figure 2.5. The flagged areas are those with no colour (i.e. white). The parameters used in the SERPent run were $(\beta_1/N_1^{max}/\beta_2/N_2^{max}/\sigma_k) = (5/2/5/2/3.0)$, where the sub-scripts 1 and 2 denote first and second runs of the algorithm.

TABLE 2.2: SERPent flagging parameters used in the treatment of RFI across all observations epochs of the COBRaS Legacy L-band data set. The parameters are given in a format of $(\beta_1/N_1^{max}/\beta_2/N_2^{max}/\sigma_k)$ where $N^{max}$ is the maximum subset size, $\beta$ is the aggressiveness parameter, $\sigma_k$ is the kick out sigma value and the subscripts denote each of the two full runs of the flagging algorithm (see Section 2.2.2 for further details).

| **Baseline** | **Target Fields** CYG-L-(A-G) | **Phase-Cal** J2007+404 | **Flux-Cal** 1331+305 | **Point-Cal** 1407+284 |
|---|---|---|---|---|
| 1-5 | 5/2/5/2/2.8 | 10/4/12/16/2.8 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 1-6 | 5/2/5/2/2.8 | 10/4/12/16/2.8 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 1-7 | 4/2/4/2/2.5 | 10/4/12/16/2.8 | 4/2/4/2/2.5 | 3/2/3/2/2.4 |
| 1-8 | 4/2/4/2/2.5 | 10/4/12/16/2.8 | 4/2/4/2/2.5 | 3/2/3/2/2.4 |
| 1-9 | 4/2/4/2/2.6 | 10/4/12/16/2.8 | 5/2/5/2/2.6 | 4/2/4/2/2.5 |
| 2-5 | 5/2/5/2/2.8 | 10/4/12/16/2.8 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 2-6 | 5/2/5/2/2.8 | 10/4/12/16/2.8 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 2-7 | 4/2/4/2/2.5 | 8/4/10/8/2.4 | 4/2/4/2/2.5 | 3/2/3/2/2.4 |
| 2-8 | 4/2/4/2/2.5 | 8/4/10/8/2.4 | 3/2/3/2/2.4 | 3/2/3/2/2.4 |
| 2-9 | 4/2/4/2/2.6 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 4/2/4/2/2.5 |
| 5-6 | 5/2/5/2/2.8 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 5-7 | 4/2/4/2/2.6 | 8/4/10/8/2.4 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 5-8 | 4/2/4/2/2.5 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 3/2/3/2/2.4 |
| 5-9 | 4/2/4/2/2.6 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 4/2/4/2/2.5 |
| 6-7 | 5/2/5/2/2.8 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 5/2/5/2/2.6 |
| 6-8 | 4/2/4/2/2.5 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 4/2/4/2/2.5 |
| 6-9 | 4/2/4/2/2.6 | 10/4/12/16/2.6 | 5/2/5/2/2.6 | 4/2/4/2/2.5 |
| 7-8 | 3/2/3/2/2.4 | 10/4/12/16/2.6 | 3/2/3/2/2.4 | 3/2/3/2/2.4 |
| 7-9 | 4/2/4/2/2.6 | 8/4/10/8/2.4 | 4/2/4/2/2.5 | 4/2/4/2/2.5 |
| 8-9 | 4/2/4/2/2.6 | 8/4/10/8/2.4 | 4/2/4/2/2.5 | 4/2/4/2/2.5 |




(A) Pre-SERPent Run.




(B) Post-SERPent run.

FIGURE 2.8: An amplitude plot of the RR polarisation, Lovell to Knockin (1-5) baseline of the phase-calibrator source, J2007+404. (A) Before running SERPent - at this stage the flagmask has been applied and large areas corrupted with RFI had been manually flagged, and (B) the data has been passed through SERPent with parameters shown in Table 2.2, yet some of the Lovell stationary scans and all the zero level dropouts have been missed, causing the SumThreshold sequence to 'overflag' good visibilities as shown in IFs 7 and 8.

trial and error process across a good range of the parameter space, the output 'SPPlot' with the applied FG table from each SERPent run were visually inspected. The best fitting parameters (shown in Table 2.2) for each baseline of every source, were those that: (i) best identified and mitigated the RFI and (ii) were most computationally efficient in terms of time[2].

When inspecting the results of the application of SERPent on the phase calibration source, J2007+404, it became evident that the flagger had not successfully identified and removed all of the Lovell stationary scans and a large number of the zero level dropouts. As a result, the statistics calculated within the SumThreshold sequence were skewed due to the increase in 'off-source' low level amplitude visibilities. This caused SERPent to flag visibilities that belonged to the true signal from the bright phase calibrator. Figure 2.8b shows a section of the RR polarisation of the Lovell to Knockin (1-5) baseline to illustrate the 'over-flagging' caused by the failure to eliminate both the zero level drop outs and the Lovell stationary scans. Both the zero level dropouts and Lovell stationary scans have been correctly identified in IF 4, yet missed in every other IF. This effect is purely down to the level of RFI present in the data set. By inspection of Figure 2.8a, a plot showing the same section data before the application of SERPent, IF 4 has the minimum level of RFI compared to its neighbouring spectral windows. Every other IF has a sufficient amount RFI that SERPent is unable to correctly identify the Lovell stationary scans, and those dropout visibilities seen at the start of each scan.

SERPent was altered in order to specifically deal with this problem. An option was included that adds together the identified zero-level dropouts and Lovell stationary scan flags from each IF to apply these over all of the spectral windows. Therefore, those correctly identified stationary scans and zero-level dropouts in IF4 were applied to the remaining IFs. In the rare circumstance where SERPent was unable to correctly identify the Lovell stationary scans and zero-level dropouts in any of the IFs, manual flagging techniques were employed. Having correctly identified and flagged the stationary scans across all of the Lovell baselines and mitigated the zero-level dropouts, the phase-calibrator was passed

[2]Much of the parameter optimisation for SERPent was carried out by a summer student, Alan Hung Shuo Su.

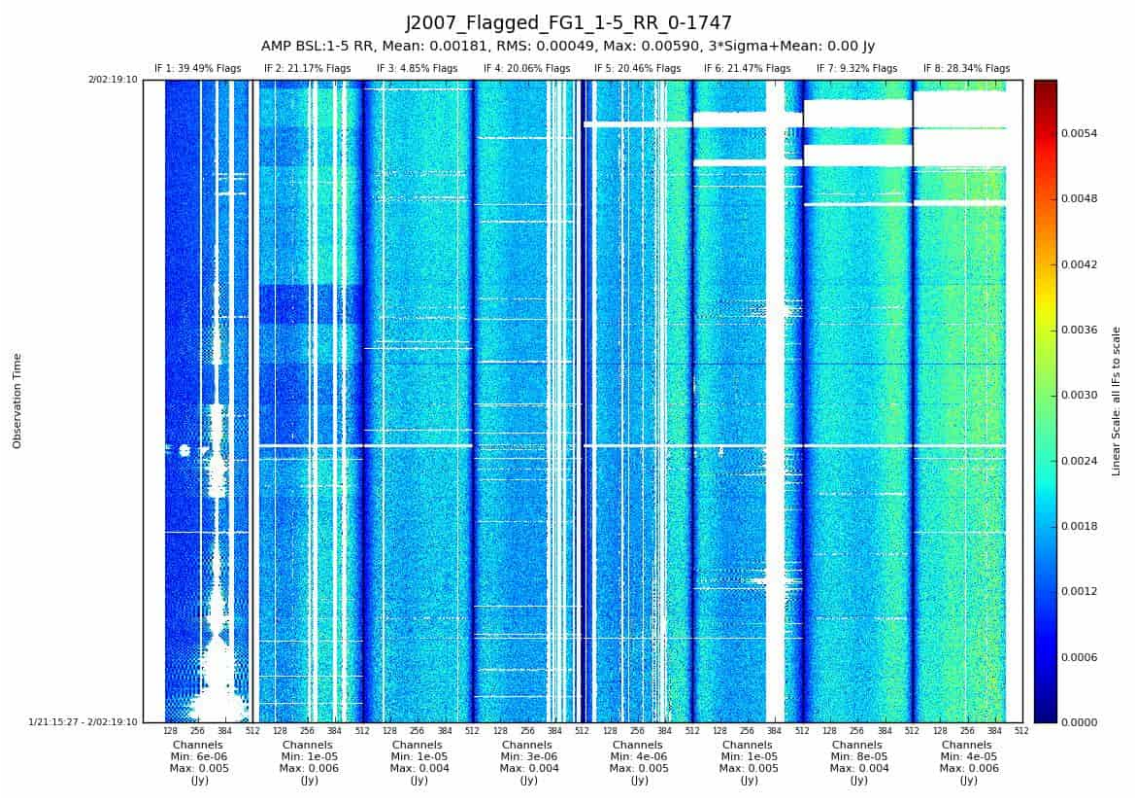



FIGURE 2.9: An amplitude plot of the RR polarisation, Lovell to Knockin (1-5) baseline of the phase-calibrator source, J2007+404. The zero level drop outs and Lovell stationary scans have been removed and both automated and manual flagging techniques have mitigated the major sources of RFI.

through the SumThreshold sequence of SERPent to good effect. Figure 2.9 shows the same section of data belonging to the phase-calibrator as shown in Figure 2.8, yet after a successful mitigation of the RFI and unusable visibilities due to in-scan dropouts and Lovell stationary scans. The amplitude scale is now a linear progression and its variation across the time and frequency domain appears significantly more uniform.

Once the entire L-band data set had been treated for RFI and unusable (off-source) visibilities, the first and last 80 channels of IFs 1 and 8 respectively were also flagged. From an inspection of Figure 2.8b, you can see that the visibility amplitudes drop off at the very start and end of the frequency band, rendering their calibration impossible. At this stage of the data reduction process, each source was ready to be calibrated. Low-level RFI became apparent throughout the various stages of calibration, since corrupt visibilities will lead to erroneous calibration solutions. Therefore, at each calibration stage (see Section 2.3), the solutions were inspected, as were the data with the solutions applied. Erroneous visibilities were correctly identified and flagged to further mitigate remaining RFI.

Table 2.3 highlights the target field and calibration sources across the two primary observation epochs contributing to the COBRaS Legacy L-band data set. Their initial data sizes and integration times (measured respectively from the size of their raw fits files and the integration time as given in their AIPS NX table) are given before any flagging had commenced. Similarly, their final data sizes (as measured from the size of their post calibration catalogue entry within AIPS[3]) and integration times (calculated relative to the amount of data lost across time and frequency space using SPPlot) are given, as is the percentage of the initial data that now remains. Across both observation epochs, $\sim$50% of the data was lost, unable to undergo a successful calibration. This is down to a variety of different factors but predominantly due to RFI and telescope errors. Other factors contributing to this loss in data include the Lovell stationary scans within the phase-calibrator and the zero-level dropouts within the scans across all sources. Furthermore, there were sections in the data where a particular antenna would 'drop-out', i.e. where a fault had occurred at the antenna such that it simply was not measuring the astronomical signal. Due to a

[3]Due to the difference in file sizes between those measured within AIPS and those in a UNIX environment, a factor was applied to account for this disparity.

TABLE 2.3: The observed sources, their data sizes and integration times across both observation epochs that contribute to the entire COBRaS L-band Legacy data set. The initial integration times and data sizes represent those of the raw data set, the last two columns however, represent the data having gone through the entire flagging and calibration process. The % of data remaining represents the amount of data left in comparison to that of the raw data files.

| Source | Initial Integration Time (hrs) | Initial Data Size (GB) | % of Data Remaining | *Final Integration Time (hrs) | Final Data Size (GB) |
|---|---|---|---|---|---|
| **2014 Apr. 11$^{th}$** | | | | | |
| CYG-L-A | 1.14 | 23 | ∼36% | 0.40 | 10 |
| CYG-L-B | 1.14 | 23 | ∼51% | 0.58 | 13 |
| CYG-L-C | 1.11 | 22 | ∼53% | 0.59 | 13 |
| CYG-L-D | 1.14 | 23 | ∼38% | 0.43 | 10 |
| CYG-L-E | 1.14 | 23 | ∼56% | 0.64 | 15 |
| CYG-L-F | 1.14 | 23 | ∼46% | 0.52 | 12 |
| CYG-L-G | 1.14 | 23 | ∼35% | 0.40 | 10 |
| J2007+404 | 3.93 | 78 | ∼40% | 1.59 | 29 |
| 1331+305 | 0.50 | 9 | ∼75% | 0.38 | 7 |
| 1407+284 | 0.50 | 9 | ∼74% | 0.37 | 7 |
| **Ap11 Total** | 12.88 | 256 | ∼46% | 5.90 | 126 |
| **2014 Apr. 26$^{th}$** | | | | | |
| CYG-L-A | 3.81 | 76 | ∼41% | 1.55 | 40 |
| CYG-L-B | 3.67 | 74 | ∼41% | 1.52 | 37 |
| CYG-L-C | 3.67 | 74 | ∼43% | 1.58 | 38 |
| CYG-L-D | 3.67 | 74 | ∼34% | 1.25 | 31 |
| CYG-L-E | 3.42 | 65 | ∼45% | 1.54 | 37 |
| CYG-L-F | 3.42 | 65 | ∼46% | 1.46 | 35 |
| CYG-L-G | 3.42 | 65 | ∼39% | 1.35 | 36 |
| J2007+404 | 8.10 | 163 | ∼37% | 2.98 | 54 |
| 1331+305 | 0.84 | 17 | ∼88% | 0.74 | 15 |
| 1407+284 | 0.99 | 20 | ∼85% | 0.84 | 17 |
| **Ap26 Total** | 35.01 | 693 | ∼42% | 14.81 | 340 |
| **Total** | 47.89 | 949 | ∼44% | 20.71 | 466 |

*This final integration time represents the relative amount of time left on source, given that the data is flagged in both the frequency and time direction.

similar amplitude level to the rest of the data, this only became evident when inspecting the phases of the phase-calibration source, J2007+404. This was the case for a large time chunk of the Apr. 26th data across all Defford (antenna #6) baselines where there was no coherent phase pattern. Another contribution to the data loss came from all visibilities associated with the LL polarisation of all Pickmere baselines. The noise within the amplitudes and phases were significantly higher for these visibilities relative to the rest. These data were calibrated but significantly added to the noise within the resulting images. Their omission on the other hand resulted in a $\sim 20\%$ increase in the sensitivity of the resulting images.

Flagging 'bad' data is absolutely necessary to obtain the highest sensitivity, since corrupted visibilities will introduce errors within the calibration and produce artefacts within the resulting images. However, having flagged such a significant portion of the COBRaS L-band data, the theoretically predicted sensitivity of the observations (see Table 2.1) will be impossible to reach.

## 2.3 Calibration of COBRaS L-band data set

At this stage of the data treatment process, the observations had been thoroughly inspected and rid of 'unusable' visibilities. Each observation epoch was processed individually. Throughout the following section (2.3), a general description of the calibration and imaging routines will be provided, whilst any problems specific to either data set will be addressed. For further information regarding any of the AIPS tasks used throughout this Chapter (and Thesis) please refer to Appendix A. Furthermore, for reference, please see Table 2.4 below regarding the names, numbers and sizes of the antennas within the e-MERLIN array.

### 2.3.1 Calibration outline

The calibration procedure was based upon the official e-MERLIN cookbook (version 3.0 - February 2015) which can be downloaded from the e-MERLIN website. The cookbook was

TABLE 2.4: The antenna elements of the e-MERLIN array.

| Antenna Number | Antenna Name | Dish Diameter (m) |
|---|---|---|
| #1 | Lovell | 76 |
| #2 | Mark II | 25 |
| #5 | Knockin | 25 |
| #6 | Defford | 25 |
| #7 | Pickmere | 25 |
| #8 | Darnhall | 25 |
| #9 | Cambridge | 32 |

written by a team based at the University of Manchester and was designed specifically for the treatment, calibration and imaging of e-MERLIN data. Several calibration pipelines exist for the treatment of e-MERLIN data (see Argo 2014 and Chapter 3 of Peck 2014). However, due to inconsistencies (and variations in data quality) between the COBRaS L-band Legacy data sets, their reduction required a bespoke treatment, whereby each step of the calibration routine could be visually inspected, with the possibility of applying further flags. Figure 2.10, a flow diagram, highlights the main steps taken from loading the raw fits files into AIPS, through to the point of self-calibration (see Section 2.3.6).

The diagram includes the initial 'data loading and house-keeping' steps used to load the data into AIPS before any flagging began. Individual AIPS tasks are represented by cyan coloured hexagons and tasks carried out by scripts written in PARSELTONGUE are represented by red rectangles. Having loaded the data into AIPS, the first AIPS task UVSORT, sorted the data into a time baseline order. INDXR, another AIPS task was used to re-index the data using the time gap between each scan to write an NX table containing the number of visibilities for each scan associated with each source, including the length, start and end time of each scan. The AIPS task UVFLG was then used to flag the Lovell to Mk2 (1-2) baseline, which as discussed in Section 1.5, is unusable due to its short baseline length in comparison to those associated with the rest of the array.

The antenna (AN) table associated with each raw fits file contained an error. This was true of all e-MERLIN data observed before January 2015. The AN table describes the relative positions and diameters of each antenna belonging to the array. The position axis offset of the Cambridge antenna was set to zero, yet its true position was offset by 1.503

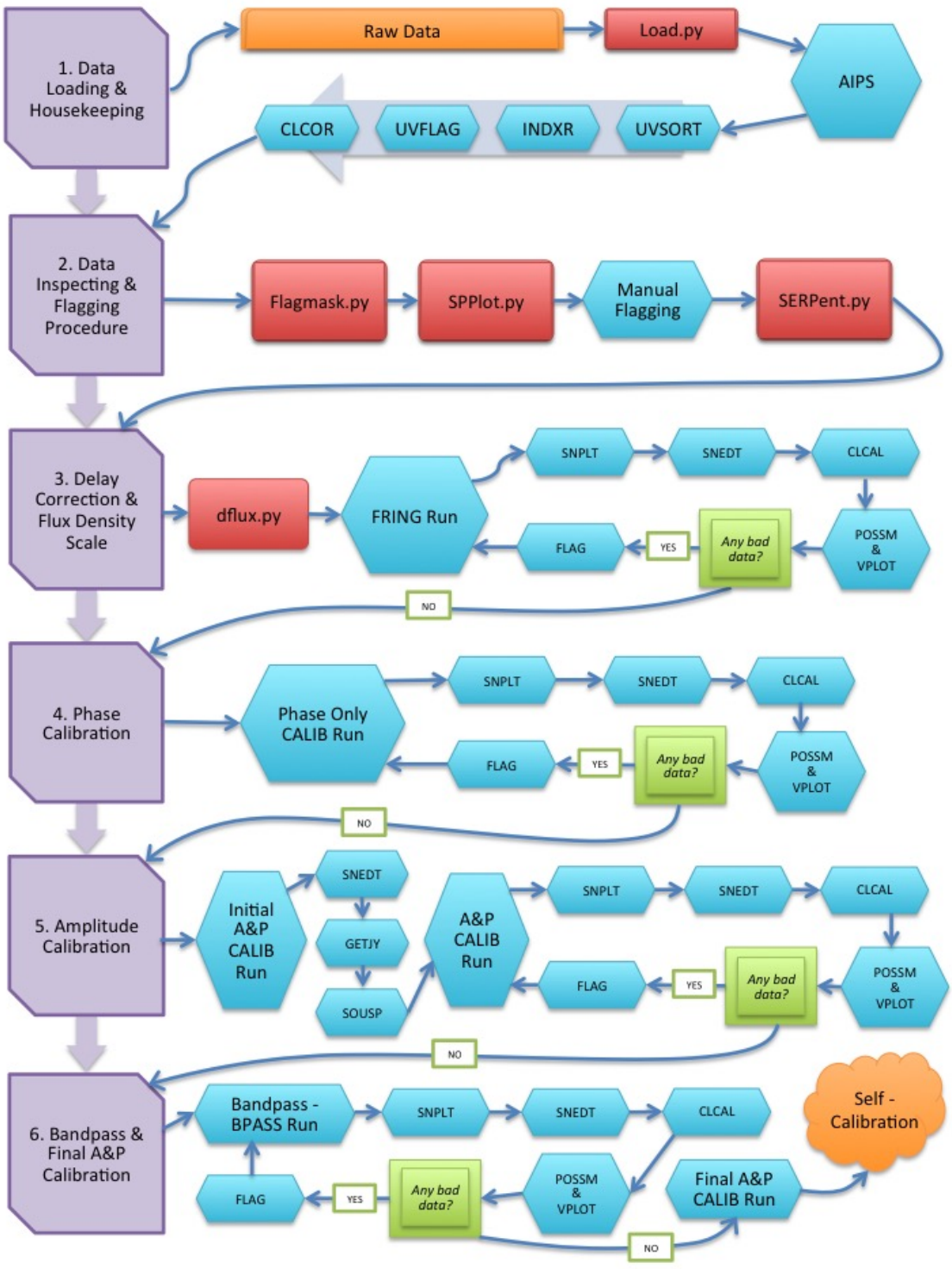



FIGURE 2.10: A flow diagram describing the steps taken in the treatment of the COBRaS Legacy L-band data from loading the raw data files into AIPS, through to the point at which the data is ready for self-calibration. Steps carried out with the use of AIPS tasks are highlighted by cyan coloured hexagons, whilst those using scripts written in PARSELTONGUE are surrounded by red rectangles.

m. This needed to be corrected before beginning the calibration, since this offset could potentially introduce serious errors into the subsequent calibration solutions. Unfortunately, this error was completely unknown until the beginning of 2015, at which point the entire data set from the Apr. 11th epoch had been calibrated. As such, the observations had to be completely re-calibrated having corrected for the Cambridge antenna position offset. Having performed the basic 'data loading and house-keeping' tasks and corrected for the Cambridge antenna position offset by use of the AIPS task CLCOR, the various flagging routines described in Section 2.2 were completed. At this stage, the data was in a position to begin the calibration routines.

### 2.3.2 Delay correction and flux density scale

To correctly calibrate the amplitudes of the observed target fields and calibration sources, the flux of at least one of the calibration sources must be known. In the case of e-MERLIN, the flux calibrator 1331+305 or otherwise known as 3C286, a bright quasar of known flux is used. At L-band, 1331+305 is heavily resolved by the longer baselines of e-MERLIN and has a steep spectral index. The fluxes must be set for each IF to account for this variation across the wide bandwidth, as shown in Table 2.5. This is accomplished using a PARSELTONGUE script called `dflux.py` that calculates the flux of 1331+305 at the central frequency of each IF. The calculation uses a polynomial expression derived from the most recent spectral flux densities of 1331+305 from Perley and Butler (2013), which is then corrected for the higher resolving power of e-MERLIN over the JVLA. The output values from the script were then entered into the source (SU) table with use of the AIPS task SETJY.

The next task is to fit and remove the delays and rates on all of the calibration sources by a process known as fringe fitting. Delay offsets within e-MERLIN data can be as large as a micro-second and are generally larger on longer baselines. They can vary in time either as a gradual drift due to changes in the temperature affecting the length of the fibres connecting the antennas, or as sharp changes originating within the correlator. As such, they have to be corrected for, as a function of time, by use of the AIPS task FRING. The

TABLE 2.5: The known flux at the central frequency of each of the 8 L-band IFs of the flux-calibrator source 1331+305 as observed with e-MERLIN. These fluxes were calculated using the shortest usable baseline length of 1100 m, connecting the Mk2 and Pickmere antennas. The spectral index, $\alpha$ of 1331+305 across the 8 IFs is also given.

| IF | Frequency (MHz) | 1331+305 Flux (Jy) |
|---|---|---|
| 1 | 1286.33741 | 15.4443 |
| 2 | 1350.33741 | 15.1064 |
| 3 | 1414.33741 | 14.7857 |
| 4 | 1478.33741 | 14.4812 |
| 5 | 1542.33741 | 14.1917 |
| 6 | 1606.33741 | 13.9159 |
| 7 | 1670.33741 | 13.6528 |
| 8 | 1734.33741 | 13.4015 |
| Spectral Index, $\alpha$ | | -0.46428 |

phase of the interferometer is a function of the frequency multiplied by the delay, meaning delay offsets will show up as a phase slope across the frequency band. Furthermore, if the delay offset changes with time, so will the phase slope, resulting in an additional change in phase as a function of time, known as the fringe rate (Cotton, 1995).

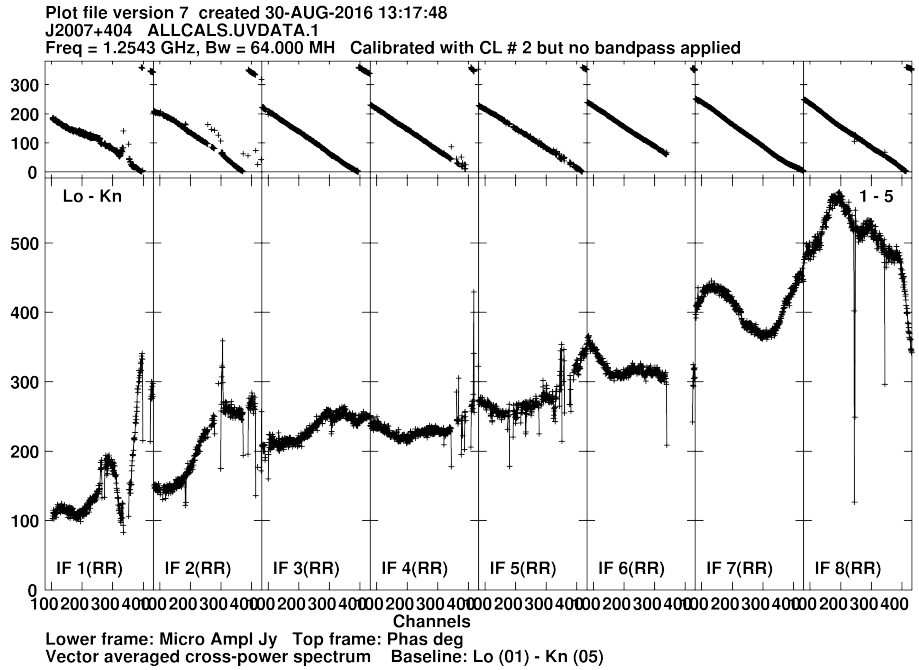
FRING finds solutions calculating a least squares fit to the phase, delay and fringe rate, storing them in a solution (SN) table attached to the AIPS datafile. These solutions, when applied to the data, remove the variable delay offsets and fringe rates. The task includes a number of pre-set parameters, the most important of which include the solution interval (`solint`) and delay window (`dparm(2)`). The solution interval is the time period for which a model is fitted to the data, i.e. the time over which FRING calculates solutions for the phase slope as a function of both time and frequency. The delay window simply defines the range of values for which to search for any delays and is given in nano-seconds (ns).

Initially, the `solint` and `dparm(2)` values were chosen to be 2 mins and 450 ns respectively for both observation epochs of the COBRaS L-band dataset. Whilst this resulted in good solutions for the Apr. 11th observations, successfully removing the phase slopes, this was not the case across the Apr. 26th observations. At certain times throughout the Apr. 26th data there existed delay jumps far larger than the 450 ns delay window given in the input parameters, meaning FRING was unable to find solutions for such large delay offsets. To correct for this, an initial FRING run with both a large delay window and solution interval

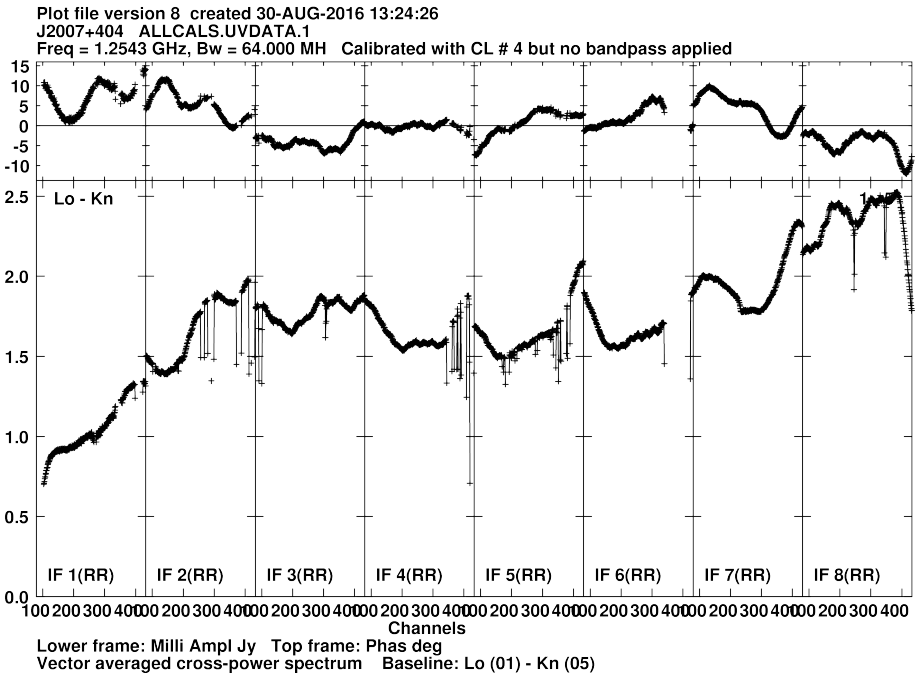
of 2000 ns and 5 mins respectively was first completed. A second run of FRING was then necessary to correct for the smaller delay jumps and the rates by using smaller solution interval of 2 mins and a delay window of 300 ns.

The delay solutions from FRING are inspected using the AIPS task SNPLT, which can be viewed on the AIPS TV or saved to a plot (PL) file. If any individual outliers exist, it is probable that they are unable to be calibrated and are removed using the AIPS task SNEDT. This task plots the solutions of the SN table in question on the AIPS TV and allows the user to interactively flag any erroneous bad solutions, creating a new SN table. Once happy with the delay solutions stored in the SN table, the task CLCAL is used to write a new calibration (CL) table having applied the solutions to a pre-existing CL table. At this point, the data are visually inspected with the newly created CL table applied, i.e. with the delay solutions applied to the data. The data is best inspected in both frequency and time, using the AIPS tasks POSSM and VPLOT to respectively do so. This inspection is crucial throughout the entire calibration process to ensure the derived solutions have been correctly implemented and the visibilities are in the desired state to move onto the next stage in the data reduction process. If the POSSM and VPLTS show any erroneous data points (or in the case of the delay solutions for example, have been unable to remove the phase slopes across the IFs), it is necessary to find the visibilities responsible, flag them, re-run FRING and go through the same inspection process.

Figure 2.11 shows a before and after POSSM plot of the RR polarisation of baseline 1-5 (Lo-Kn), having run FRING twice on the phase-calibrator of the Apr. 26th dataset. In Figure 2.11a, the phase slopes across each IF due to the uncorrected delay offsets are clearly visible in the upper panel. Having corrected for these delay offsets as a function of time and frequency, Figure 2.11b now shows that the phase (upper panel) variation is restricted to a range of $\sim 15°$, with these fluctuations likely due to variable atmospheric delays which are corrected for in the subsequent calibration. Having checked every baseline and polarisation of each individual calibration source (in this case 1331+305, 1407+284 and J2007+404) to ensure the delay offsets have been correctly removed, the delay correction step is complete.




(A) Pre-FRING.




(B) Post-FRING.

FIGURE 2.11: Amplitude and phase plots created using the AIPS task POSSM, showing the RR polarisation of baseline 1-5 (Lo-Kn) of the phase-calibrator J2007+404 with (A) showing the data before any correction of the delay and rates and (B) after two FRING runs in order to correct for the large delay jumps found across some of the baseline. These observations were part of the Apr. 26th epoch of the COBRaS L-band Legacy dataset. These POSSM plots have been averaged over time a period of $\sim$ 4.3 hours, accounting to roughly half of the phase-cal data from the Apr. 26th data set.

### 2.3.3 Phase calibration

The next stage in the calibration routine involves running the AIPS task CALIB in order to compute antenna based gain solutions by initially solving for the phases only. The phase solutions are calculated across each of the three calibration sources, however only the solutions gained from the phase-calibrator, J2007+404 are used to calibrate the phases of the target fields. J2007+404 was chosen as a phase calibrator source as it is bright, lies within close proximity (within $\sim 2°$) of our target fields and even on the longest e-MERLIN baselines at L-band, is not resolved enough to introduce significant phase errors.

A bright, point source is required because in Fourier space, this represents a constant phase, making phase errors easy to identify and therefore solutions easy to obtain. Furthermore, since phase errors are introduced due to atmospheric disturbances, the phase calibrator is chosen to be as close to the target field as possible to ensure the phase solutions represent the same pathway between the target field and the interferometer. The phase-calibrator is observed periodically between target field scans over a time length chosen to be less than the coherence time. The coherence time describes the length of time over which the electromagnetic wave at our observing frequency can be considered predictable despite phase instabilities due to atmospheric effects within the ionosphere and troposphere. At L-band (21 cm), if considering the longest Mk2 - Cambridge baseline ($\sim$ 217 km), the coherence time is $\sim$ 40 minutes in good conditions (Peck, 2014). However, to minimise the risk, the phase calibrator is sampled approximately every 7.5 mins ($\sim$ 15 mins for all Lovell baselines) to ensure the phases stay coherent.

CALIB is initially run on the flux calibrator, 1331+305 which is resolved on longer e-MERLIN baselines and requires an input model to aid its calibration. Since the phases have been known to vary rapidly (on timescales less than a minute), an initial solution interval of 0.5 mins is chosen for the phase-only CALIB runs. The first run outputs the solutions into a new SN table. The outputs from the second CALIB run on the point source and phase calibrator 1407+284 and J2007+404 respectively, are written into the same SN table. At the end of each CALIB run, the AIPS message window displays the number of solutions that were found and that failed. If the number of failed solutions are greater

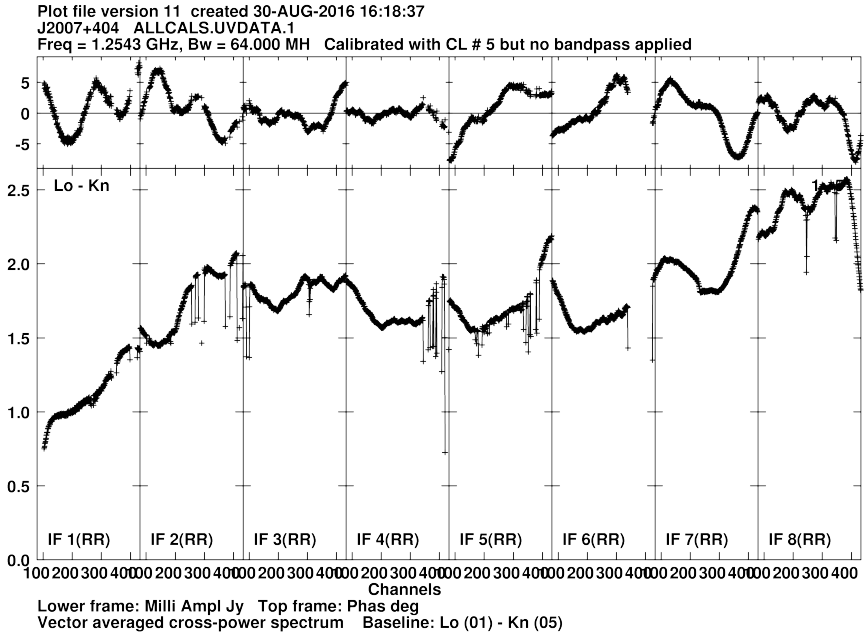
than $\sim 5\%$ of those found, there is likely some remaining corrupt visibilities present within the data set, whose identification may not become apparent until the solutions have been inspected or even applied to the data set.

As with the solutions from FRING, the phase-only solutions from CALIB are first inspected using SNPLT. The presence of individual outliers are flagged using SNEDT. Once satisfied with the phase-only solutions, CLCAL can again be used to interpolate these solutions with those written in the previous CL table. POSSM and VPLOT are used to plot the data in both the frequency and time domain with the newly found phase solutions for each of the three calibration sources. Examples of outputs from both plots of the phase-calibrator, J2007+404 can be seen in Figure 2.12. At this stage, it may become apparent that some of the phase solutions were wrong if any of the averages exhibit significant fluctuations or phase shifts. By identifying the baseline, IF, channel or time at which the error exists, these visibilities can be flagged manually using such AIPS task as SPFLG, IBLED or UVFLG (for further details on each of these AIPS task, refer to the Appendix A). An example can be seen in the VPLOT shown in Figure 2.12b around a time of 2/00-2/01 hours where some high phase values exist. In this case, the points exhibiting phase values larger than 50 degrees within that time range were flagged.

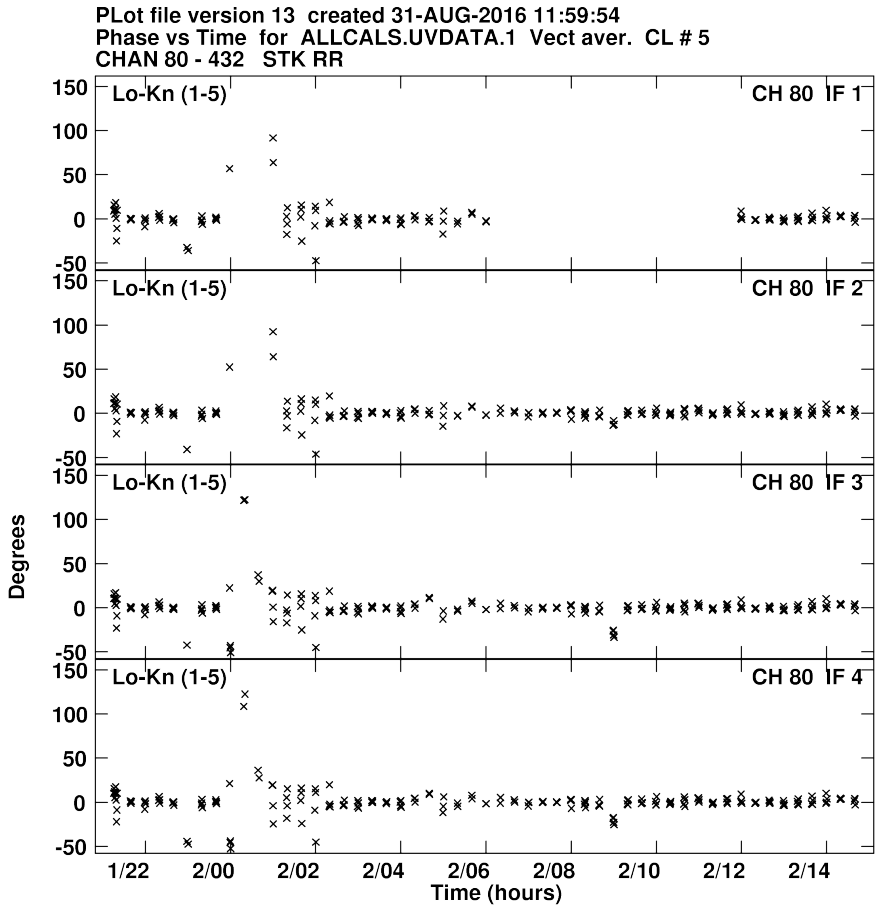
### 2.3.4 Amplitude calibration

This section of the calibration routine aims to set the absolute flux level for the point source, phase calibrator and target fields. The measured amplitudes of each of the observed sources outputted from the correlator are set to an arbitrary level. The absolute flux level in Janskys (Jy) of the other calibration sources and the target fields can be set by using the well known flux of 1331+305, and the scaling of the arbitrary units between each of these sources.

This is put into practice by completing two initial runs of CALIB to solve for both the amplitudes and phases of firstly the flux-calibrator 1331+305, and secondly of the point source and phase calibrators 1407+284 and J2007+404 respectively. The results of both of these initial CALIB runs are saved into the same SN table which is subsequently inspected




(A) POSSM output - Frequency vs Phase (upper panel) and Amplitude (lower panel); averaged in time.




(B) VPLOT output - Time vs Phase for IFs 1 to 4; averaged in frequency.

FIGURE 2.12: (A) POSSM plot and (B) VPLOT of the RR polarisation of the 1-5 baseline of the phase calibrator, J2007+404, having been passed through a phase-only run of CALIB to correct for the atmospheric induced phase errors. Both plots represent the same chunk of data as shown in Figure 2.11.

with SNPLT. This SN table containing the amplitude solutions for each of the calibration sources is used to set the absolute flux scale of 1407+284, J2007+404 and the target fields by using the AIPS task GETJY. In order to obtain reliable results with GETJY, the amplitude solutions within the SN table must be completely clean of outliers. SNEDT is used to flag these outliers and any other solutions that do not appear to fit the general trend. These initial amplitude solutions are purely used to set the fluxes, and the SN table is deleted having done so. As a result, it is best to use sections of the data that look the cleanest.

TABLE 2.6: The final derived fluxes of the point and phase calibration sources of the COBRaS L-band Legacy data set. These values were derived using the Apr. 26th observations and were applied to those from Apr. 11th. The final values used are those under the SOUSP column in either case.

| **IF** | **Frequency** | **1407+284** | | **J2007+404** | |
|---|---|---|---|---|---|
| | (GHz) | GETJY (Jy) | SOUSP (Jy) | GETJY (Jy) | SOUSP (Jy) |
| 1 | 1.28633741 | 0.7696 | 0.7701 | 4.5802 | 2.2050 |
| 2 | 1.35033741 | 0.8289 | 0.8233 | 2.2548 | 2.2443 |
| 3 | 1.41433741 | 0.8842 | 0.8774 | 2.2702 | 2.2834 |
| 4 | 1.47833741 | 0.9349 | 0.9324 | 2.3259 | 2.3193 |
| 5 | 1.54233741 | 0.9570 | 0.9883 | 2.3430 | 2.3553 |
| 6 | 1.60633741 | 1.0500 | 1.0449 | 2.3743 | 2.3902 |
| 7 | 1.67033741 | 1.1122 | 1.1024 | 2.4355 | 2.4243 |
| 8 | 1.73433741 | 1.1612 | 1.1606 | 2.4652 | 2.4575 |
| Spectral Index, $\alpha$ | | 1.34317 | | 0.35499 | |

Having edited the initial amplitude solutions, creating a new SN table, GETJY is run on both 1407+284 and J2007+404 in order to derive their absolute flux values across each IF and saves them to the source (SU) table. Furthermore, the AIPS task SOUSP is used to fit a straight line to the derived flux values over each IF to determine the spectral indices of the point source and phase calibrator. The derived spectral flux densities for 1407+284 and J2007+404 from the Apr. 26th observations can be seen in Table 2.6. For J2007+404, GETJY derived a flux in IF1 of $\sim$ 4 Jy, about a factor of 2 higher than the value derived in every other IF. This highlighted that IF1 was still likely affected by RFI and had little data remaining in comparison to the other spectral windows. Before running SOUSP, the flux of IF1 was set to $\sim$ 2.2 Jy such that the polynomial fit from SOUSP was not skewed by the inaccuracy of the original flux density derived in IF1.

The Apr. 26th observations constituted a much larger portion of the entire Legacy data

set and were of a better quality than those of the Apr. 11th. As a result, these flux values were deemed more reliable than those derived from the Apr. 11th dataset, which were within 10% of those shown in Table 2.6. Moreover, these fluxes were manually applied to the Apr. 11th data set using the AIPS task SETJY to achieve a good consistency between the two observation epochs.

In setting the fluxes of the point source and phase calibrators using a combination of GETJY and SOUSP, the SN tables containing the amplitude solutions could be deleted. CALIB was run in exactly the same way again, solving for the amplitude and phases for the flux, point source and phase calibration sources. This time however, utilising the derived absolute flux values of 1407+284 and J2007+404 in the SU table to set the correct flux scale in the resulting SN table from CALIB. In exactly the same process as before having run FRING and a phase-only CALIB run, the resulting SN table was inspected with SNPLT, but this time both phase and amplitude solutions were inspected. Similarly, SNEDT was used to flag any individual outliers before running CLCAL to apply the solutions to the previous CL table containing both the phase only and fringe fitting solutions. Again the data were inspected across both time and frequency using VPLOT and POSSM tasks.

If the amplitude and phase solutions from CALIB were successful, the phases across time and frequency should be further restricted to within $\sim 10^{\circ}$ and the vector averaged amplitude for both 1407+284 and J2007+404 should equal their respective absolute flux values found in the SU table. Figure 2.13 shows the same chunk of data as before, the RR polarisation of the 1-5 baseline for J2007+404, yet with the amplitude calibration having been applied to the data. The variation in the phases and amplitudes over the frequency band are now purely down to the response of the receiver. This is corrected for in the next section, through the bandpass calibration.

### 2.3.5 Bandpass calibration

The AIPS task BPASS corrects for the response of the interferometers receivers, which cause complex amplitude and phase gain variations across the frequency band. In addition, the bandpass calibration must also account for the spectral index and curvature intrinsically

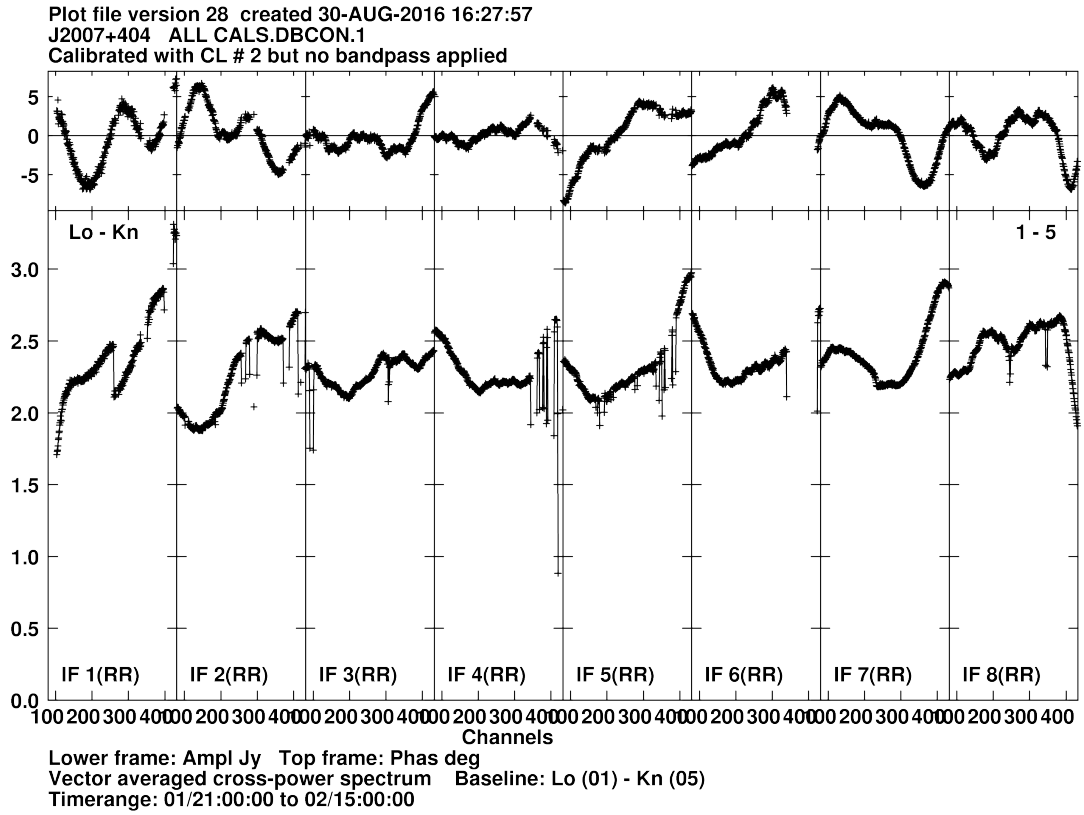



FIGURE 2.13: A POSSM plot of the RR polarisation of the 1-5 baseline of the phase calibrator, J2007+404, having been passed through an amplitude and phase run of CALIB to apply the derived fluxes found with GETJY and SOUSP shown in Table 2.6. This plot represents the same chunk of data as shown in Figures 2.11 and 2.12.

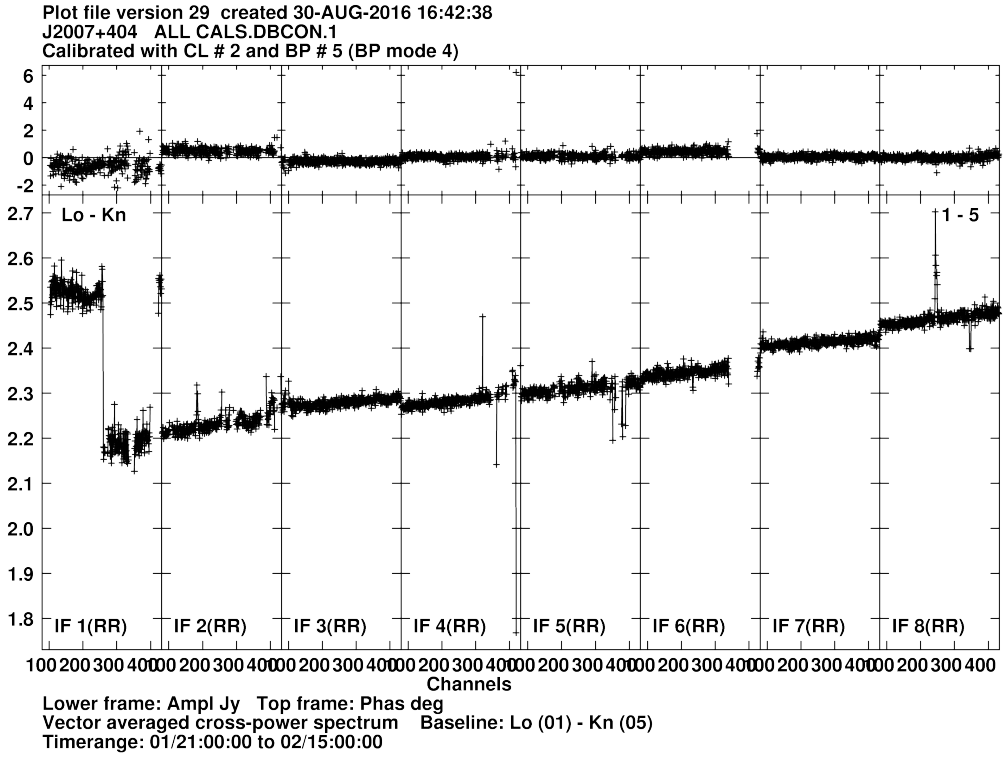
associated with each source. In order to solve for the variations across the frequency band, BPASS uses a point source calibrator that is both bright and has a flat spectrum across the band. In the COBRaS observations, this source is 1407+284, (otherwise known as OQ208), however the phase-calibrator, J2007+404 can also be used as the bandpass calibrator if need be.

1407+284 was used as the initial bandpass calibrator to find solutions for the variation in amplitude and phase across the frequency band. The bandpass solutions calculated from 1407+284 of the Apr. 26th observations did not correctly remove the amplitude and phase variations across the entire frequency band for a number of baselines when applied to the phase-calibrator. Clearly, the bandpass response varied as a function of time, and as a result the small time range sampled by the observations of 1407+284 were not sufficient to correct for the bandpass response over the entire time range sampled by the target field observations. Therefore the phase-calibrator J2007+404 was used to create bandpass solutions over each scan to account for the variable response across the frequency space. This produced much better results, flattening the amplitudes and phases across each IF, as can be seen in Figure 2.14a.

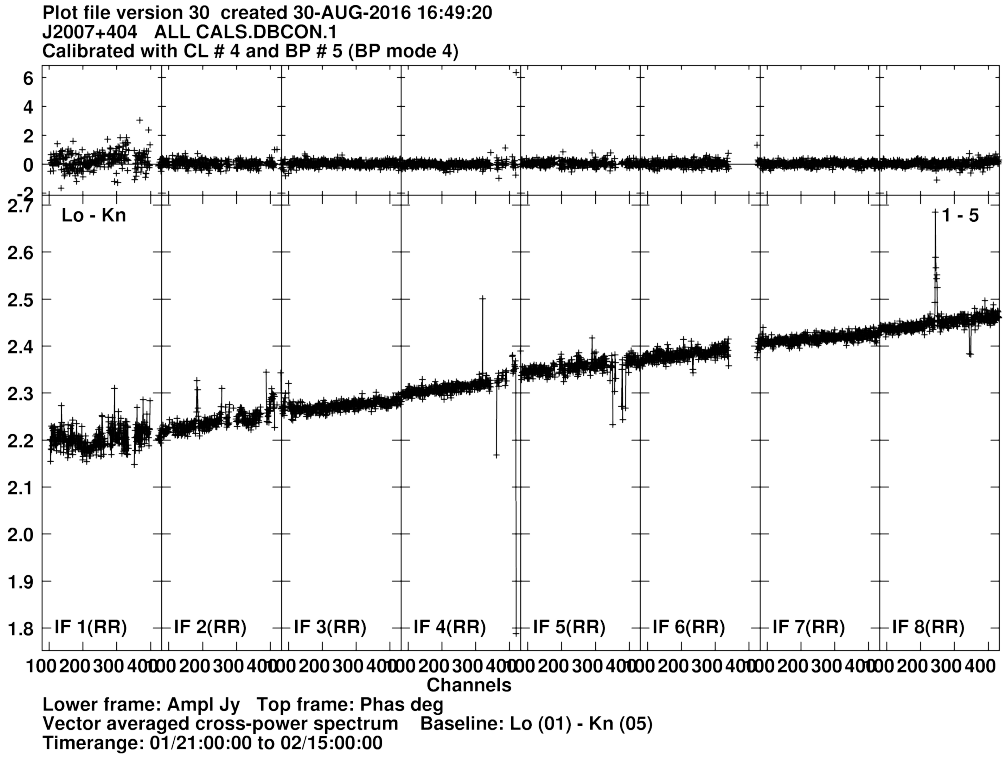
In the application of the calculated bandpass solutions (Figure 2.14a), the phases and amplitudes have been flattened within each IF, however a final amplitude and phase CALIB run is necessary to re-correct both the phase and amplitudes with the bandpass solutions applied. This ensures that the amplitude and phases line up from IF to IF, as can be seen in Figure 2.14b. This final amplitude and phase CALIB run was run solely on the phase-calibrator, and the same steps were taken in inspecting and editing the solutions with SNPLT and SNEDT. In applying the solutions to the most recent CL table with CLCAL, the POSSM and VPLOTs were inspected and any 'bad' data was flagged.

### 2.3.6 Self-calibration

Self-calibration is a process whereby the remaining phase or amplitude errors within the visibility data can be corrected for. The routine involves running the AIPS task CALIB to calculate additional phase and amplitude solutions, which are determined using the




(A) Bandpass applied.




(B) Bandpass and final A&P CALIB applied.

FIGURE 2.14: Two POSSM plots to show the effect of the applied bandpass on the RR polarisation, baseline 1-5 of the phase-calibrator J2007+404. (A) shows the effect of the bandpass being applied and (B) shows the application of the final CALIB run solving for the amplitudes and phases having previously applied the bandpass calibration. Again, this is the same chunk of data as those shown in Figures 2.11, 2.12 and 2.13.

clean components (CCs) obtained from a previous model of one (or more) bright source(s) within the data. A completely point-like source is not essential, yet it must have a sufficient signal-to-noise ratio to ensure accurate phase solutions can be calculated within CALIB from the CCs produced with IMAGR. IMAGR is the primary AIPS task that deals with the deconvolution and imaging of radio visibilities (see Section 2.4.1 for further details).

Before the self-calibration procedure can begin, the calibration and bandpass tables obtained through the previous calibration steps must be applied to the target fields as well as the other calibration sources (where appropriate). In regards to the COBRaS Legacy L-band data sets, the previous calibration steps were completed upon a single AIPS catalogue entry that included all three of the calibration sources. This was to minimise the size of the working data file in order to minimise the time spent running each AIPS task throughout the calibration routine. At this stage, the target fields now needed to be combined with the calibration sources such that the AIPS task CLCAL could be run on each of the relevant SN tables, applying the previous calibration steps to the target fields. Furthermore, the AIPS task CLIP was run on each of the target fields and phase calibrator as a function of baseline, in order to remove obvious outliers (large amplitude visibilities) still present within the data.

At this point, the self-calibration routine as depicted within the flow diagram of Figure 2.15, was ready to be implemented. The strategy is to first self-calibrate the phase-calibrator J2007+404, because its flux density is high enough to obtain good incremental corrections, which can then be applied to the target fields as was done with the previous calibration solutions. For both observation epochs of the COBRaS L-band Legacy data, an initial image of the phase-calibrator was produced. An example of this initial image of J2007+404, taken from the Apr. 26th observations, can be seen in Figure 2.16 with a peak flux of ∼2.348 Jy and a Root Mean Square (RMS) noise level within the image of $\sim$ 202 $\mu$Jy, giving a signal-to-noise ratio, $S/N \approx 12000$. For comparison, the initial image of J2007+404 produced at the same calibration stage for the Apr. 11th data gave a peak flux of 2.344 Jy, RMS = 744 $\mu$Jy, and a signal-to-noise of $\approx$ 3150.

Following the flow diagram shown in Figure 2.15, the number of clean components is

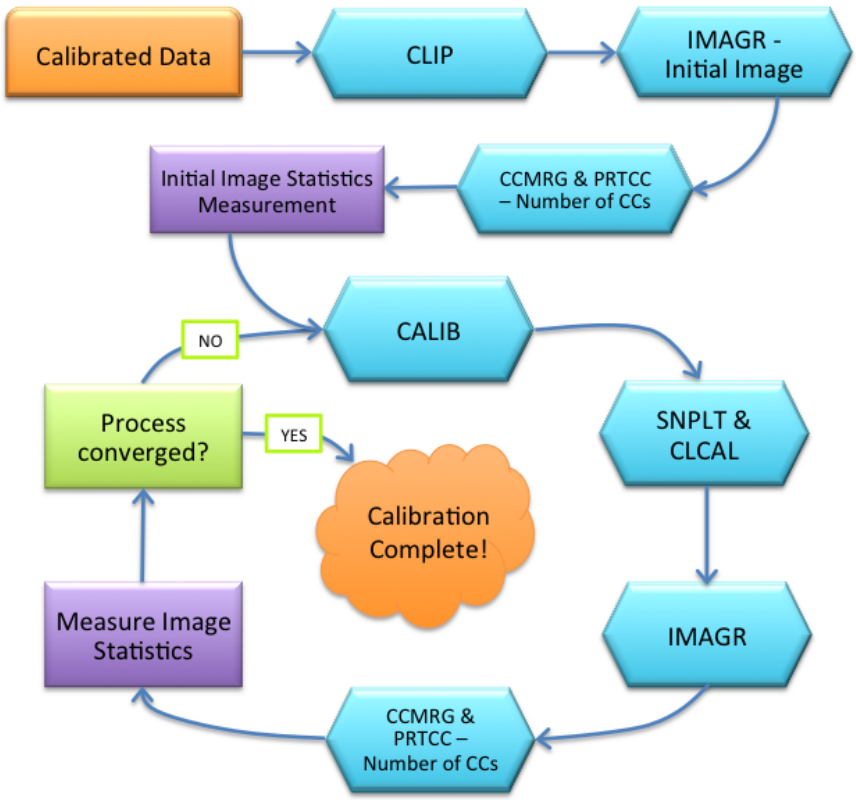



FIGURE 2.15: A flow diagram highlighting the steps taken throughout the self-calibration process of the COBRaS L-band Legacy target field data. AIPS tasks are highlighted by cyan coloured hexagons and tasks completed manually such as measuring the image RMS and peak on source flux, are represented by purple rectangles.

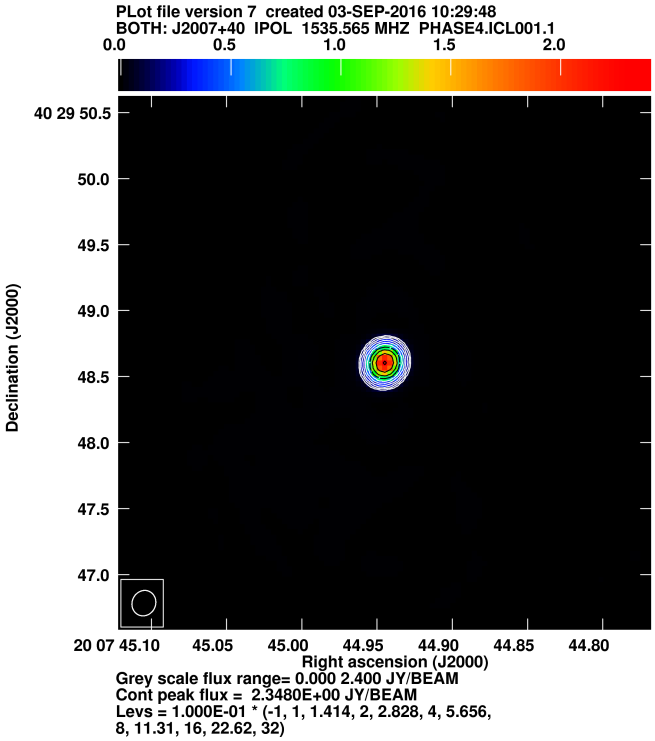
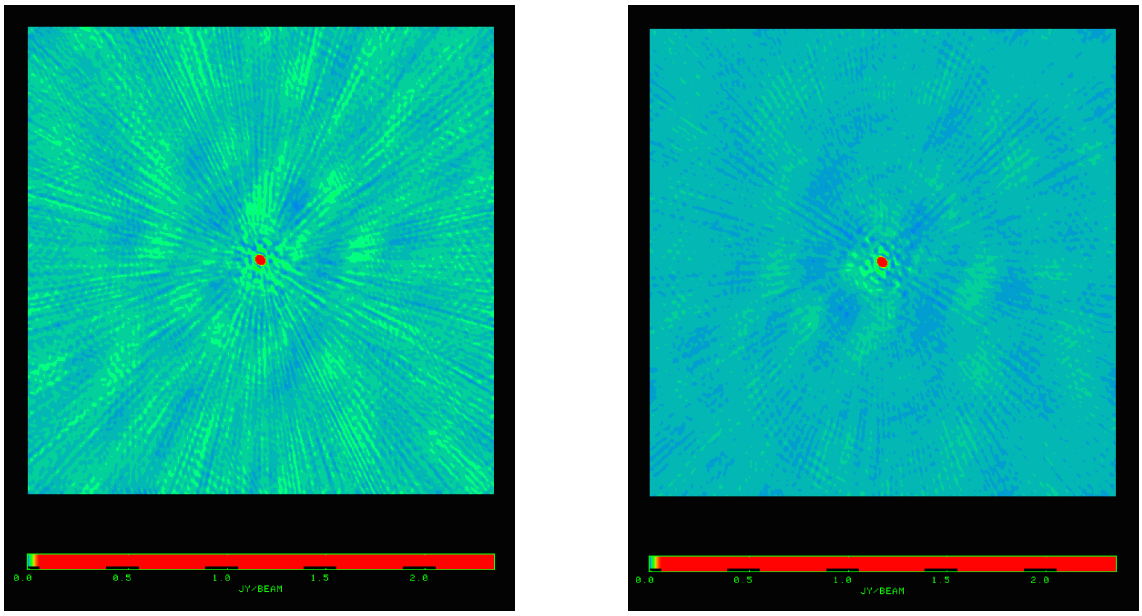



FIGURE 2.16: An initial image of the phase-calibrator, J2007+404 before any self-calibration attempts. This image was produced from the COBRaS L-band Legacy observations taken on 2014 Apr. 26th. Integrated flux = 2.351±0.235 Jy, image RMS = 202 $\mu$Jy.

obtained using a combination of the AIPS tasks CCMRG and PRTCC to first merge, then print the CCs on the terminal window. The number of CCs to use within the following CALIB run must be chosen carefully. As a general rule of thumb, the number of CCs is chosen up to that which presents the first negative flux within the image, since the following CCs are likely to not be associated with the source. Artefacts within the image, or perhaps the resolved nature of the object in question, may also be present in the CCs past the first negative. These CCs will not represent the phase of the bright calibration source, meaning CALIB will obtain poor solutions if such CCs are included. Before running CALIB, the peak and RMS of the image are measured both on- and off-source. These measurements, along with the general state of the image itself, serve as a reference for comparison with the following images, produced through successive iterations of the self-calibration cycle.

Throughout the self-calibration, CALIB solves for the phases only, outputting the solutions into a new SN table. As usual, these can be plotted and edited using SNPLT and SNEDT respectively before CLCAL is used to interpolate the solutions with the previous CL table. Note that at this stage the phase solutions should not exceed 20° and should evolve consistently and smoothly as a function of time. The next image is produced with IMAGR and the same steps are taken to obtain the number of believable CCs and image statistics. The image RMS, peak on- and off-source flux and the produced solutions are compared to those found from the previous iteration. Whilst the overall image signal-to-noise should increase, the variation in the phase solutions should decrease. The entire self-calibration process will have converged when i) the peak on-source value no longer increases ii) the phase solutions vary by less than 1 degree and/or are dominated by noise and iii) the image looks indistinguishable from the last.

After six iterations of self-calibration upon the phase calibrator of the Apr. 26th observations, the signal-to-noise of the successive images were approximately equal to that of the initial image. Additionally, the phase solutions of the original calibration (prior to any self-calibration) were constrained to $\sim 0.5°$ (with the exception of a few outliers which were subsequently flagged). After six iterations of self-calibration the phase solutions were constrained no further and it was deemed that self-calibration was unable to improve the original phase solutions. As a result, no self-calibration was applied to the phase-calibrator




(A) Initial: pre self-calibration (B) Final: post self-calibration

FIGURE 2.17: A comparison between two images of the phase calibrator, J2007+404 taken from the COBRaS L-band Legacy observations taken on Apr. 11th. (A) The initial noisy image prior to any self-calibration attempts; peak on-source flux = 2.3435, RMS = 744 $\mu$Jy. (B) The final image having undergone six rounds of self-calibration; peak on-source flux = 2.3519, RMS = 433 $\mu$Jy.

for the Apr. 26th observations. The artefacts present and poor signal-to-noise within the initial image of J2007+404 of the Apr. 11th observations however, meant that significant improvements could be made by self-calibration. Figure 2.17 shows the initial image of J2007+404 in comparison to the final image having undergone six rounds of self-calibration. The radial artefacts present in the initial (left) image are mostly absent in the final (right) image. Moreover, after six iterations the peak on-source flux has increased whilst the image RMS has decreased marking an improvement in the signal-to-noise by nearly a factor of two, from $\sim$ 3150 to $\sim$ 5430.

Self-calibration was also applied to each of the seven target field pointings from both the Apr. 26th and 11th datasets. This was accomplished by first producing an initial, low resolution, wide field image of each of the target fields. From this image, the brightest sources ($\gtrsim$ 1 mJy) within each field were chosen to perform the self-calibration process with. Only the brightest sources were chosen as significant flux was required in order to calculate reliable phase solutions with CALIB. Despite using $\sim$ six sources per field, CALIB

failed to find good phase solutions. To remedy this, solutions were calculated over a larger time interval (10 mins) and by first averaging together the RR and LL polarisations. The application of these phase solutions to the data, resulted in a lowering of the RMS noise level whilst increasing the peak value of the sources within each field.

Figure 2.18 shows one example of the effect that each successive iteration of the self-calibration has on one of the sources found in pointing B of the Apr. 26th observations. Moving from left to right, each successive plot shows how some of the flux around the immediate vicinity of the source moves directly on top of it, increasing the peak on-source flux and signal-to-noise ratio. By the fourth self-calibration iteration, the process had converged since there is little difference between bottom left and right images shown in Figure 2.18. This process was carried out for each of the seven target field pointings across both of the observation epochs contributing to the COBRaS L-band Legacy data. The final CL tables were applied to the data, signalling the end of the entire calibration routine.

## 2.4 Imaging of the COBRaS Legacy L-band data

The majority of the imaging procedures were undertaken within AIPS, primarily by the use of the task IMAGR. The following section (2.4.1) will give a general overview of the imaging process whilst discussing how IMAGR makes use of the CLEAN algorithm in order to deconvolve the dirty image with the synthesised beam. Section 2.4 will also go on to explain how imaging can be a powerful tool in diagnosing errors that may remain within the data due to low level RFI or calibration errors. Furthermore, Section 2.4.3 will explain how the entire wide field maps of each of the target field pointings were produced.

### 2.4.1 Imaging procedure and the clean algorithm

In the completion of the calibration routine as described in Section 2.3, the data was ready for deconvolution and imaging. This process, completed by the AIPS task IMAGR, utilises the CLEAN algorithm to deconvolve the dirty image, i.e. the radio visibilities sampled over points in the $u, v$ plane, with a synthesised beam or point spread function (PSF), in order

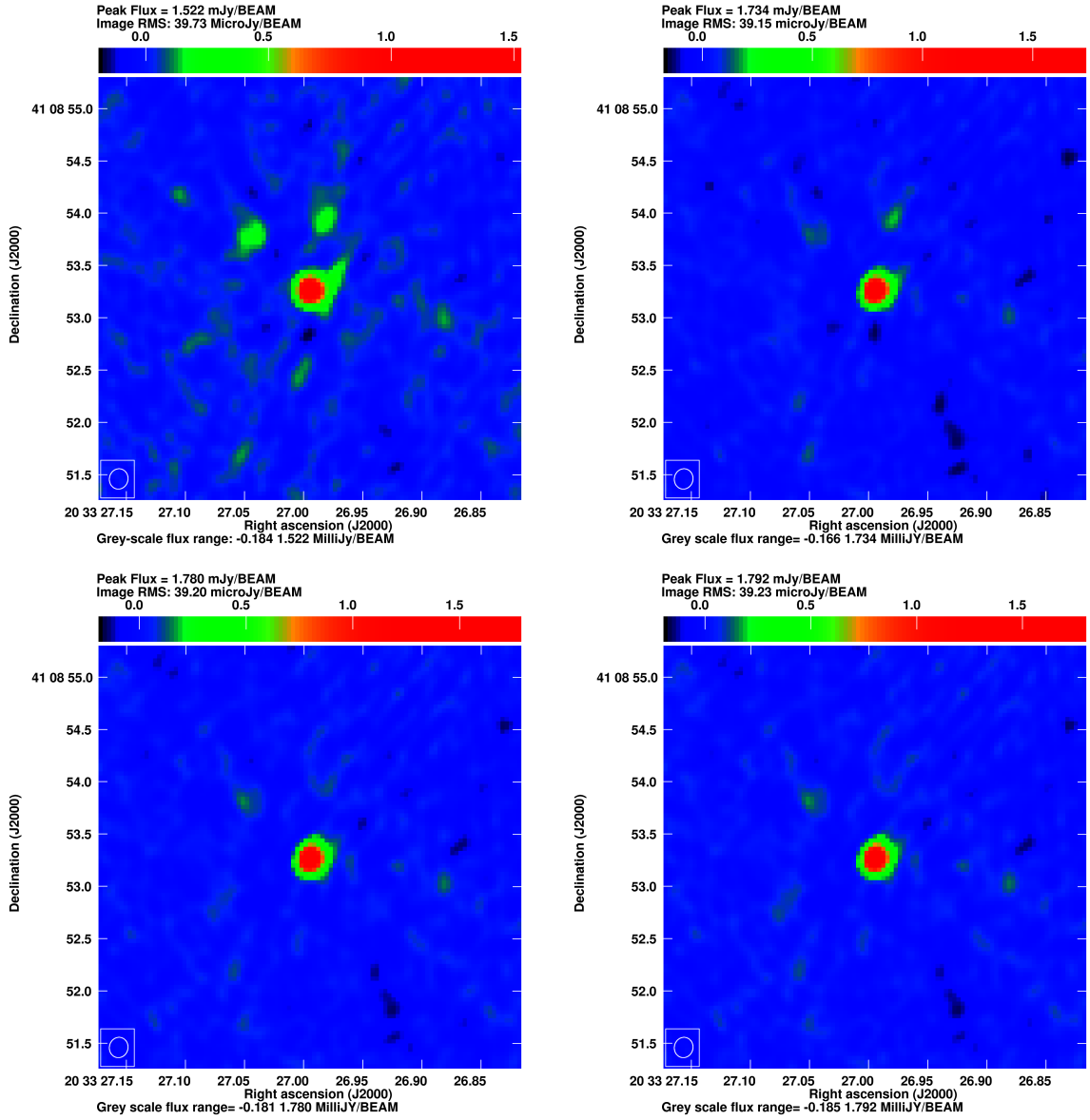



FIGURE 2.18: The successive images produced from the self-calibration routine of pointing B, the target field from the COBRaS Legacy observation taken on Apr. 26th. The initial image is shown at the top left, the first iteration top right, third on the bottom left and the final iteration on the bottom right.

to obtain the true image of the sky (see Equation 2.14). For IMAGR to carry out this deconvolution process, both the measured visibilities, $V_\nu(u, v)$ and the sampling function $S(u, v)$ must undergo a Fourier transform to obtain the dirty image and PSF. This is accomplished by distributing the data over a grid from which Fast Fourier Transforms (FFTs) are performed on the $u, v$ data. By gridding the visibilities in this way, different weighting schemes can be employed upon the PSF to alter the properties of the final image. The different weighting schemes, which are specified within IMAGR and applied to the grid, can either emphasise or 'smooth-out' particular visibilities, depending on their position in the $u, v$ plane.

Natural weighting, gives an equal weight to each visibility, meaning gridded cells containing a larger concentration of visibilities are emphasised, increasing the image sensitivity at the cost of decreasing its resolution. Uniform weighting on the other hand applies an equal weighting to all the grid cells, effectively giving more weight to the longer baselines, increasing the resolution at the cost of decreased sensitivity. Robust weighting offers a middle ground between natural and uniform weighting that can be tweaked by the use of a single parameter.

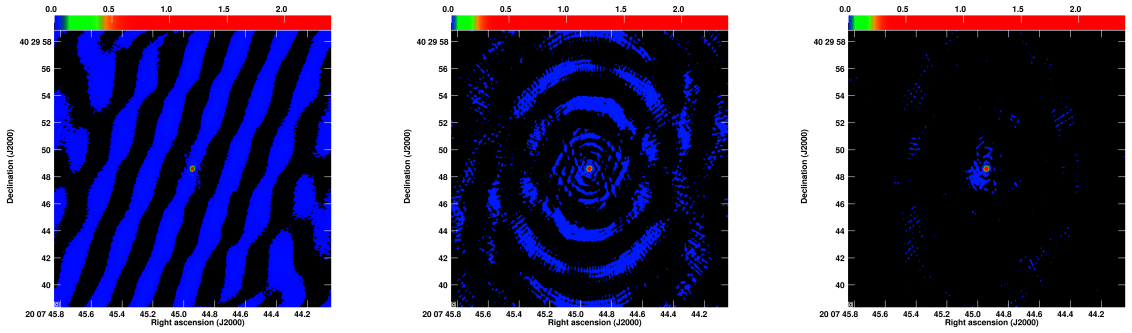
The original CLEAN algorithm was developed by Högbom (1974) and assumes that all radio sources, or positions of flux within the data can be represented by point sources. An iterative procedure firstly identifies the position and strength of the peak, found within the dirty image. This peak is multiplied by some gain factor (typically $\leq 1$), convolved with the dirty beam (or point spread function) and the result of which is then subtracted from the dirty image. The position and strength of the subtracted peak (i.e. a clean component - CC) is recorded in a model as a delta function. These steps are iterated until no more peaks above a particular threshold level exist within the dirty image, which is now referred to as the residual image. The point source clean components are then restored with a clean beam, normally represented by a Gaussian function in order to de-emphasise the higher spatial frequencies (i.e. those represented by the longer baselines) which are often spurious. The restored point source model is then added to the residual map to form the final clean or true sky map.

The deconvolution algorithm implemented within IMAGR uses a re-worked CLEAN algorithm known as the Cotton-Schwab CLEAN algorithm (Schwab, 1984; Cornwell, Braun, and Briggs, 1999). This algorithm stemmed from another version by Clark (1980) that fundamentally differs from the original by the use of 2-dimensional FFTs. The successive CLEAN algorithms further differ by iteratively finding the positions and strengths of the point source components using only a small patch of the dirty beam known as a beam patch. This results in an algorithm that is both computationally more efficient and sufficient in obtaining CCs for PSFs with large side-lobe patterns.

### 2.4.2 Initial images and artefact hunting

Initial images of the phase-calibrator and target fields were produced both before and throughout their respective self-calibration procedures (see Section 2.3.6). These initial images were used to check the general artefact level and the sensitivity of the data, whilst also identifying and imaging the bright sources that went into building a model for self-calibration. They covered $\sim$ 20" squared on the sky and were produced using a natural weighting scheme. Each had an image size of 512 $\times$ 512 pixels and a cell size of 0.04" per pixel corresponding to the full resolution of e-MERLIN at L-band. Firstly, the images of the phase-calibrator, J2007+404 could be used to measure the success of the calibration. As a point source, the resulting images should show a bright source in the centre of the field of an amplitude equal to that derived in the amplitude calibration. There should exist no strong artefacts within the surrounding field, which can manifest themselves as radial and non-radial striations of varying low level amplitudes.

Figure 2.19 shows three initial images of the phase-calibrator highlighting some examples of the artefacts present in the data. The first image (Figure 2.19a) shows some strong diagonal striations across the image likely due to some remaining phase errors present within one of the baselines. To further narrow down the error, the AIPS task FFT was used to Fast Fourier Transform the produced image back into the $u, v$ plane. This will highlight the position of the error causing the dominant diagonal striations in the $u, v$ plane, allowing one to diagnose which baseline is responsible. Sharp peaks in one domain




(A) (B) (C)

FIGURE 2.19: Three images of the phase calibrator, J2007+404 from the COBRaS L-band Legacy observations taken on Apr. 26th. The contrast has been altered to highlight the background artefact level present in each image, with the level decreasing in the successive images moving from left to right. Each image is of size 512 × 512 pixels with a cell size = 0.04". The CLEAN algorithm was run with 500 iterations in each case with a natural weighting scheme.

(e.g. in the $u, v$ plane), will transform to a broad feature in the other (e.g. the image plane). The broad diagonal striations shown in Figure 2.19a resulted in a peak in the $u, v$ plane associated with one of the inner baselines. Each of these baselines were inspected in time and frequency space as a function of both phase and amplitude with the task SPFLG. Some spurious phases for a number of scans associated with IFs one and two of the 7-8 (Pi-Da) baseline (which also manifested as amplitudes larger than the average values) were attributed to the artefacts shown in Figure 2.19a. Their removal from the data resulted in the image shown in Figure 2.19b. Error recognition in the image plane can provide a powerful tool to diagnose remaining errors in the $u, v$ visibilities.

Figure 2.19b shows no non-radial striations, however the general artefact level appears quite noisy with some radial artefacts present, representing the shape of the dirty beam. This indicates some general noisy amplitudes within the data which are eliminated via the AIPS task CLIP. This task flags all visibilities with amplitudes above a given threshold level. By plotting the amplitudes (in Janskys) as a function of $u, v$ distance for each antenna (using the AIPS task UVPLT), one can determine the threshold level to clip as a function of antenna. In this instance, (i.e. for the phase calibrator data from the Apr. 26th observations), threshold values of 3 Jy and 10 Jy for the Lovell and each of the other six antennas repetitively resulted in a reduction in the noise level from $\sim$ 574 $\mu$Jy to $\sim$

202 $\mu$Jy. The image resulting from the application of CLIP is shown in the third and final image (Figure 2.19c) which appears relatively artefact free. The production of such an image indicates that the calibration has been successful and the determined phase and amplitude solutions appear to accurately represent those of a point source, i.e. those of J2007+404.

A further test of the reliability of the calibration solutions occurs throughout the self-calibration routine. The iterative process should obtain phase solutions constrained to within a smaller range, whilst the resulting RMS and peak on-source flux within the images should decrease and increase respectively until convergence is reached (see Section 2.3.6 for further details). However, as was the case for the Apr. 26th epoch of the COBRaS L-band observations, several rounds of self-calibration on the phase-calibrator J2007+404 produced phase solutions and images of no better quality than after the initial calibration routine. Whilst this indicates a reliable calibration of the phase-calibrator, calibration errors may still become apparent when applying the calibration to the target fields and inspecting their initial images.

Initial images of the target fields were made of a low resolution, imaging a large proportion of the entire field of view, prior to any self-calibration in order to check the general artefact level, the presence of which needed to be eliminated as they can affect the subsequent self-calibration of the target field data and therefore the reliability of the resulting full-resolution maps. Figure 2.20, shows an example of one of these initial wide field images from the Apr. 26th observations of pointing E. Some strong radial artefacts can be seen centred just to the right of the image centre, as well as some diagonal striations across the entire image. The radial artefacts which dominate the image are due to bad visibilities consistent across all frequencies. They originated from the beginning of each scan where the antennas were likely still off-source, i.e. zero level dropouts. However, the automatic flagger SERPent was unable to detect them. This is because the COBRaS target pointings contain no bright sources and therefore the average amplitude across all the visibilities was similar on-source and off-source. The AIPS task QUACK provides a simple way to flag a set period of time from either the beginning or end of each target scan. The length of time required to be taken from each scan ($\sim$ 10 secs) was determined by a trial and error

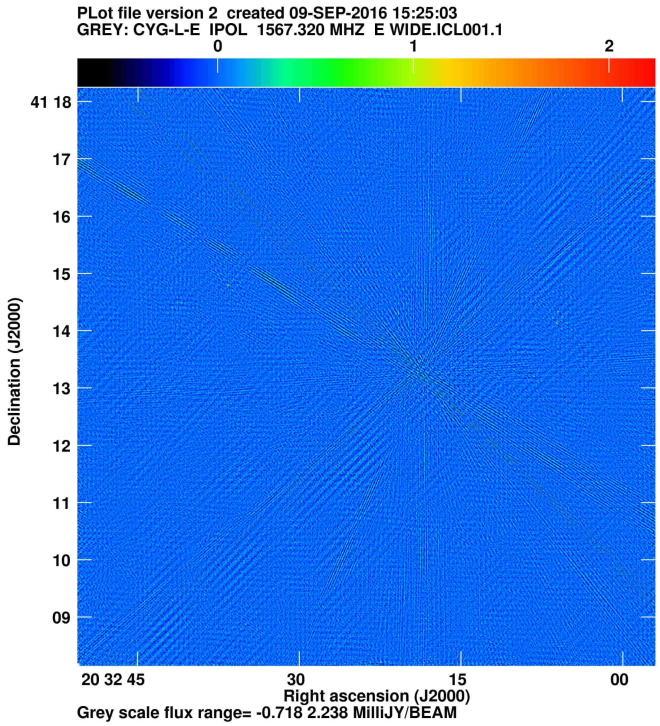
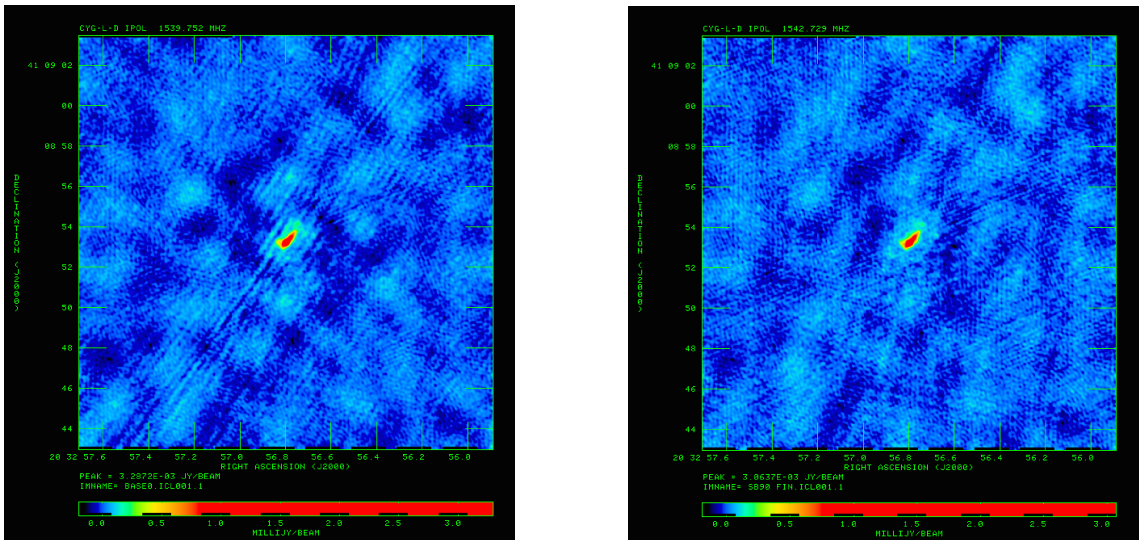



FIGURE 2.20: An initial 1024 × 1024 image of the target field pointing E of the COBRaS L-band observations taken on Apr. 26th. This is a low resolution image (cell size = 0.296"), spanning a 300" squared portion of the sky.




(A) (B)

FIGURE 2.21: Two images of SHBW90 taken from pointing D of the COBRaS L-band Legacy observations taken on Apr. 26th. These two images highlight the difference in the level of artefacts present before (A) and after (B) having removed the visibilities responsible.

process.

In satisfying oneself with the level of image artefacts across low resolution wide field images of the target, the bright sources within each field could be selected and imaged at the full resolving power of e-MERLIN at 21cm. These images were to further check for any artefacts (and therefore errors within the data) whilst also providing the initial images ready for the target's self-calibration procedure. Bright sources were chosen as self-calibration requires a suitable amount of flux in order to calculate accurate phase solutions. Moreover, multiplicative errors (i.e. those that become convolved with the source distribution) become more apparent when imaging bright sources. An example can be seen in Figure 2.21 where both images represent the currently unclassified bright source known as SHBW90 found in pointing D of the COBRaS L-band observations. This source was first identified in the Westerbork Synthesis Radio Telescope 1400 and 350 MHz Continuum Survey of Cyg OB2 Setia Gunawan et al. (2003).

The initial image shown in Figure 2.21a contains a number of different artefacts associated with some remaining errors within the data. At this stage of the data reduction process, all obvious errors in the $u, v$ plane (i.e. in phase or amplitude across both time and frequency

space) had been eliminated. Thus in order to narrow down the position of these errors within the $u, v$ data, a trial and error process was required to re-image the source without the data from a given antenna, baseline, time range or even IF. By flagging each antenna in turn and imaging SBHW90, it was possible to determine which antenna was responsible for a given artefact (if any). By then flagging each baseline in turn (associated with the responsible antenna) and imaging, the error was narrowed down to the baseline level. The data was further imaged in time chunks and the particular times (scans) and baselines held accountable for image artefacts were flagged. The majority of artefacts (such as those shown in Figure 2.21a) were associated with Cambridge baselines and in some cases with the 1-5 (Lo-Kn) and 1-6 (Lo-De). Their emission from the data resulted in a cleaner image as shown in Figure 2.21b. It is likely that such artefacts were due to remaining phase errors since they were associated with only the longest baselines in the array. This image error analysis was completed on all seven of the target fields of the COBRaS L-band Legacy data.

### 2.4.3 Production of the COBRaS wide-field L-band maps

In the completion of the target self-calibration routine and the removal of all remaining errors from the image plane, each target field pointing was ready to be wide field imaged at the full resolution available with e-MERLIN. The produced wide field images would provide the base maps from which to produce the entire source population catalogue of the COBRaS L-band Legacy data set.

Before any of the final wide field imaging began it was necessary to re-weight the data to account for the differing antenna sizes across the e-MERLIN array. The Lovell telescope, with a diameter of 76m for example, is $\sim 3$ times larger than most of the other antennas in the array. Therefore the visibilities associated with the Lovell baselines need to be weighted up to match the relative collection area of the Lovell antenna. This has the effect of increasing the sensitivity of the resulting images by adding weight to the more sensitive baselines. It can be applied using two different AIPS tasks, namely WTMOD or REWAY. Where WTMOD requires the user to specify the antennas weights, REWAY modifies the

$u, v$ weights by calculating the uncertainty (or noise) within selected portions of the data. Whilst REWAY is relatively untested on e-MERLIN data, both tasks were implemented and the resulting images in either case compared. Across the seven target field pointings, WTMOD gave the best and most consistent sensitivity in the resulting images and became the preferred re-weighting scheme throughout the COBRaS L-band data reduction process.

Initially, the wide field images were produced using the wide field routine[4] associated with the official e-MERLIN pipeline (Argo, 2014). The routine uses the AIPS task UVFIX, which shifts the phase centre of the $u, v$ data to a particular point on the sky in order to split the data into a chessboard like ($8 \times 8$) grid of smaller $u, v$ data files. Each of these 64 $u, v$ data files (covering the entire field of view of the observations) are then imaged separately. The resulting images are then combined together using the AIPS task FLATN to produce the final wide field image. This routine ensures that regions far from the pointing centre (i.e. within the side-lobes of the PSF) are not subject to imaging errors when deconvolving the dirty image with the PSF. In order for the routine to run on appropriate timescales, the data is averaged in both time and frequency when the data is split into 64 separate $u, v$ data files. However, this averaging introduced errors into the resulting wide field image, producing a large amount of noise around the sources within the field and increasing their flux densities by $\sim 30 - 40\%$ in comparison to individually produced, single, 512×512 images of each source. As a result the wide field images produced by the routine associated with the e-MERLIN pipeline were deemed unreliable and unfit for source extraction methods.

To remedy this, each pointing was first imported from AIPS, into the Common Astronomy Software Applications (CASA) package. The expected field of view of each pointing (based upon the size of the Lovell antenna) was $\sim 657$" at the central frequency of the observations. In order to avoid imaging errors within this field of view, the CLEAN task within CASA was used to image an area on the sky roughly twice as large as this ($\sim 30000$ pixels$^2$ at a cell size of $\sim 0.04$" pixel$^{-1}$). This was to avoid any potential computational errors within the imaging process by ensuring that the entire field of view (as determined by the Lovell antenna) was processed by the CLEAN algorithm. An area almost twice as large

[4]Written by Dr. Nick Wrigley based at the University of Manchester.

as this field of view was necessary since the CLEAN algorithm utilised by IMAGR will typically only use the inner portion of the image. The final area on the sky (per pointing) considered for source extraction was the central region of these images, corresponding to the field of the view of the Lovell antenna. Loading the images back into AIPS, the task SUBIM was used to create a sub-image of the central 791" × 791" region, equal to the primary field of the view of the Lovell antenna at the very top of the observing band (the highest frequency). This would ensure the maximum amount of sky was covered by the subsequent source extraction process, without any degradation due to the introduction of image errors far from the phase centre.

The wide field images produced within CASA gave a good consistency between the source flux densities of those previously imaged on an individual level with the AIPS task IMAGR. However, the noise level at a given point in the field was inherently larger than that of an individual small-field image if imaged at the same position on the sky. This difference varied as a function of distance from the pointing centre. At the centre of the field (i.e. at the pointing centre), the sensitivity of the wide field images were a factor of $\sim 1.2$ larger than that measured at the same point on the sky but from a 512×512 pixel (small-field) image. At the outer extremities of the wide field image (away from the pointing centre), this factor increased to $\sim 1.5$. Despite this slight increase in the noise level, these wide field images were deemed fit for use upon the various source extraction methods (see Chapter 3). Note that for pointing G, the close proximity of the incredibly bright micro-quaser Cyg X-3 ($\sim$ 100 mJy at cm wavelengths in its 'normal' quiescent level Waltman et al. 1994) introduced major errors into the imaging process within CASA and the program was unable to image the desired 1200" squared field of view. As a result a field of view according to the central frequency of the band ($\sim$ 657" squared) was only possible for pointing G.

The final step in the creation of the wide field maps was to correct for the response of the primary beam by dividing the image by the primary beam gain factor. This is accomplished using the AIPS task PBCOR which takes user defined parameters of the primary beam of e-MERLIN to make such a correction to the image. This results in lowering the sensitivity in areas further out from the centre of primary beam in conjunction with the true response

of each antenna.

## 2.5 Conclusions and future work

The complete reduction of the COBRaS L-band Legacy data as described above was completed by the Summer of 2016. Such a large amount of time was taken to complete the reduction due to the sheer volume of data and vast amount of RFI present across the observations, coupled with the need for a reliable and robust calibration. The key science goals of the COBRaS project demand such a vigorous application of the data reduction process in order to produce 21cm radio maps of unrivalled sensitivity, allowing for the detection of massive stars, YSOs and other stellar exotica. Moreover, the COBRaS L-band data set represents some of the initial Legacy science observations made by e-MERLIN at 21cm, meaning not only that issues were still present in the precise treatment of the data, but poor data quality and inconsistencies between data sets were a common issue. Many elements of the data reduction process highlighted above were pioneering in their application to e-MERLIN L-band data and therefore did not come without their issues. Work into RFI mitigation techniques and wide field imaging routines for example, are still on-going and will undoubtedly be improved upon.

### 2.5.1 A summary of the L-band data reduction process

The steps listed below highlight and summarise the main steps taken along the data reduction process. Starting from the initial raw data files and ending with the final COBRaS 21cm radio maps of the Cygnus OB2 association, each of the two observation epochs were passed through this reduction routine separately.

1. The raw data was first loaded into AIPS where various house-keeping tasks were completed to ready the data for calibration.

2. Before the calibration, the data was inspected and flagged in order to remove unusable data. This was accomplished by the use of SERPent, SPPlot and various manual

flagging techniques to rid the data of poor visibilities that were either corrupted by RFI or were 'off-source'.

3. The calibration routine began as described in Figure 2.10.

4. At the end of the calibration routine, images of phase calibrator were produced to provide an initial assessment of the calibrated dataset, checking for remaining errors within the data and flagging them appropriately.

5. Self-calibration of phase calibrator, J2007+404 was attempted and if successful, applied (with the rest of the calibration solutions) to the target field data.

6. Initial low resolution wide field images of the each of the target fields were produced to again test the reliability of the calibration and search for remaining errors within the target field data.

7. Bright sources within each of the target field pointings were chosen to use as part of the target self-calibration routine.

8. Having self-calibrated each of the target fields and applied the solutions, each pointing was ready to be wide field imaged at the full resolving of e-MERLIN at L-band.

### 2.5.2 Assessment of the data calibration process

Quantitatively, an assessment of the data can be made by comparing the sensitivities of the produced radio maps with those expected from the observations, given their integration time. Table 2.7 displays the final integration time of the phase-calibrator and each target field pointing across both observation epochs having been completely reduced. The column titled '% data remaining' represents the amount of data that remains for a given integration time, taking into account the number of flags that exist within the data (i.e. in both time and frequency space). The relative integration time is therefore equal to the integration time multiplied by the percent of data remaining. For the observations taken on the Apr. 26th, the average relative integration time across the seven pointings is 1.46 hours.

Considering the expected sensitivity level, given 5 hours of integration time ($\sigma = 10$ $\mu$Jy), the predicted sensitivity of 1.46 hours of data is calculated using the expression:

$$\frac{\sigma_1}{\sigma_2} = \sqrt{\frac{t_2}{t_1}} \tag{2.19}$$

where $t_2 = 5$ hrs, $t_1 = 1.46$ hrs and $\sigma_2 = 10$ $\mu$Jy. It follows that the expected sensitivity $\sigma_1 = 18.5$ $\mu$Jy. Listed in Table 2.7, the average sensitivity reached across the seven target field points is 21.7 $\mu$Jy, i.e. within $\sim 17\%$ of the expected sensitivity. Note that these calculations do not take into account the amount of data lost (due to flagging) as a function of antenna. For example the Lovell telescope is the main contributor to the expected sensitivity level due to its size in comparison to the rest of the array. Table 2.7 also shows the resolutions (beam size) obtained by imaging each field. The beam size is calculated within IMAGR and depends upon the final $u, v$ coverage of the observations. The expected resolution of e-MERLIN images at this frequency is 183 mas, which again is in good agreement with the resolutions stated in Table 2.7, with the exception of a few of the target field pointings from the Apr. 11th data set. In these cases, the amount of flagging has sometimes altered the $u, v$ coverage and care must be taken if commenting on source morphologies in the case of an elongated beam shape. Note that due to the low declination of the Cyg OB2 association, a circular beam is not expected even with a more complete $u, v$ coverage.

The resolutions, artefact levels and sensitivities reached in comparison to those expected, indicates that the calibration was successful. The reason why these COBRaS observations could not reach their predicted sensitivity level is simply due to the tremendous amount of data that has been lost through RFI and other various instrumental errors both internally (electronics, correlator) and externally (antenna pointing errors).

### 2.5.3 Future work

The COBRaS observations taken on Apr. 11th and Apr. 26th of 2014 constitute the total awarded data allocation at L-band. An additional five hours of observations were

TABLE 2.7: The final statistics of the fully calibrated COBRaS 21cm datasets. The integration times listed were taken from the final NX table within AIPS having re-run the AIPS task INDXR. The percentage of data remaining shows the amount of data left taking into account the number of flags in time and frequency space still present within the data set. The resolutions (beam size) and sensitivities were obtained on a 512×512 image (cell size = 0.04") of the centre of each field, produced using IMAGR with a natural weighting scheme.

| Source Name | Integration Time (hrs) | % Data Remaining | Beam Size (mas) | Sensitivity ($\mu$Jy) |
|---|---|---|---|---|
| **2014 Apr. 11$^{th}$** | | | | |
| J2007+404 | 2.58 | 61.9% | 205x167 | 435 |
| CYG-L-A | 0.73 | 55.3% | 456x189 | 47 |
| CYG-L-B | 0.96 | 60.7% | 192x161 | 42 |
| CYG-L-C | 0.96 | 61.0% | 194x161 | 43 |
| CYG-L-D | 0.75 | 57.6% | 434x226 | 41 |
| CYG-L-E | 1.10 | 58.0% | 221x180 | 32 |
| CYG-L-F | 0.89 | 58.4% | 258x207 | 37 |
| CYG-L-G | 0.72 | 56.2% | 353x260 | 44 |
| **2014 Apr. 26$^{th}$** | | | | |
| J2007+404 | 4.71 | 63.3% | 196x178 | 202 |
| CYG-L-A | 3.05 | 50.9% | 205x180 | 21 |
| CYG-L-B | 2.90 | 52.5% | 205x172 | 21 |
| CYG-L-C | 2.97 | 53.2% | 210x185 | 21 |
| CYG-L-D | 2.42 | 51.5% | 205x157 | 24 |
| CYG-L-E | 2.90 | 53.0% | 218x166 | 20 |
| CYG-L-F | 2.75 | 53.0% | 199x163 | 22 |
| CYG-L-G | 2.79 | 48.6% | 214x162 | 23 |

obtained on the 24th of January 2014 that have yet to be completely reduced due to their comparatively small amount of time spent on-source and a reduced data quality compared to their other epochs. However, its successful reduction will provide another observational epoch, at least for some of the brightest sources within the COBRaS L-band target fields. Furthermore, the $u, v$ data from each of the three observation epochs will be co-added together from which single L-band maps can be forged of the entire dataset, increasing the overall sensitivity of the resulting images.

The predominant issue throughout the data reduction process has been the mitigation and treatment of Radio Frequency Interference (RFI). It is accountable for a large fraction of the data lost, meaning the final time spent on-source for each of the target fields across both of the observation epochs was only $\sim 1.5$ hrs compared to the proposed 5 hrs, resulting in a lowering of the sensitivity of the radio maps by a factor of $\sim \sqrt{3.3}$. The only true solution would be to clear the frequency space from interference, yet in reality this is impossible without changing the location of the e-MERLIN array to somewhere very remote. To better mitigate the RFI, improved techniques must be implemented at both the pre- and post- correlation stages. SERPent for example will benefit from an option to alter the flagging parameters as a function of IFs, since IFs 1 and 2 were consistently shown to harbour significant amounts of RFI in comparison to others. If pre-correlation mitigation procedures could be implemented on the e-MERLIN array this could potentially hugely reduce the amount of RFI within the data, since corrupted signals from one antenna will not then be correlated with 'clean' signals from another.

A full study of all the sources of RFI across the L-band spectrum of the e-MERLIN array is necessary. This should be calculated as a function of all areas on sky across all hours of the day, throughout the course of an entire year. As such RFI could be more readily mitigated at the pre-correlation stage, providing a good model of the known RFI were readily available for a particular observation epoch. Furthermore, the known strong sources of RFI would then be less likely to 'bleed' or 'spill' into neighbouring channels leading to corruption on such large scales across the frequency space.

Furthermore, for a better mitigation and improved knowledge of the RFI, the e-MERLIN

array must become consistent in its data quality. With reference to Table 2.7 the sensitivity of the target field images of the Apr. 26th observations far outweigh those of the Apr. 11th, relative to the amount of time spent on-source. This is a result of inconsistent data quality, where the general noise within the Apr. 11th data is larger than that of the Apr. 26th. Moreover, all visibilities associated with the LL polarisation of Pickmere baselines were completely disregarded since their contribution significantly increased the noise within the resulting images. Such 'mechanical' issues must be rectified in order to deliver observations of sensitivity and resolution true to their potential.

# Chapter 3

# The COBRaS L-band all source catalogue

This chapter presents the entire population of sources detected within the COBRaS L-band Legacy radio maps. Proceeding the complete reduction and imaging of the e-MERLIN L-band observations as described in Chapter 2, this chapter begins by first depicting the tools used to provide reliable source detections across the seven target field pointings. In doing so, the Source Extraction Algorithm for COBRaS (SEAC) is presented and compared to existing source detection methods. Following a complete source extraction of the L-band maps including accurate positional information and integrated 21cm flux densities, the produced COBRaS L-band catalogue will be cross-correlated with existing catalogues of the Cygnus OB2 association. The cross-correlation process will separate out previously detected and identified sources whilst also providing a list of previously unidentified sources. In the characterisation of as many sources as possible, some fundamental analysis will be completed. The complete L-band source population catalogue presented here will form the basis of the entire COBRaS catalogue with this chapter acting as a guide for the future treatment of the COBRaS C-band (6cm) radio maps.

## 3.1 Source detection and flux extraction

In the advent of large observational surveys containing hundreds and thousands (and in the case of Gaia even billions) of sources, the need for efficient, robust and reliable source

detection algorithms has become ever more important. The challenge is to automate the detection process with the aim of minimising the number of false detections. Currently many different methods exist, most of which are tailored towards the specific characteristics of the image in question such as the types of sources present and the waveband at which the observations were taken (see Masias et al. 2012 for a recent review).

In regards to the detection of sources within radio maps, recent work for the Evolutionary Map of the Universe (EMU) project that uses the Australian Square Kilometre Array Pathfinder (ASKAP) instrument produced two source detection codes. AEGEAN (Hancock et al., 2012) and BLOBCAT (Hales et al., 2012) both utilise the floodfill algorithm (see Section 3.1.1.1) in order to first detect the sources within the field. AEGEAN assumes each source is point like and can therefore be aptly represented by a Gaussian point spread function which is subsequently fitted to each detected source to obtain its flux. BLOBCAT on the other hand performs both Guassian and non-Gaussain routines to fit each detected source and estimate its flux, allowing for a more accurate representation of resolved sources. The reliability and success rates of the floodfill algorithm implemented in either code made it a desirable choice for its application upon the COBRaS L-band radio maps.

Broadly speaking, source finding within radio astronomy involves finding a collection of pixels, each of which contain information regarding the radio flux of an astronomical source, and the shape of which contains morphological information of this source. The problem lies in being able to disentangle the surrounding background noise within the image from the pixels that actually contain source information. More specifically in regards to the observations made for the COBRaS project, many of the sources will be of a similar level to that of the surrounding noise meaning the boundary between real astronomical source and noise will often be blurred. Furthermore, the sizes, intensities and shapes of the astronomical sources present within the COBRaS field of view will undoubtedly vary dramatically, from resolved background radio galaxies to un-resolved, point-like stellar objects. This calls for a bespoke source extraction code that can return reliable source positions and fluxes for a wide range of different source types. Such a program is presented below in Section 3.1.1.

### 3.1.1 Source Extraction Algorithm for COBRaS - SEAC

SEAC was based upon an original script written by Luke Peck (see Peck 2014) who intended its use upon the fully reduced wide field radio maps from the COBRaS observations. In implementing the floodfill algorithm within the PARSELTONGUE/PYTHON environment, the original version of the code was tested on synthesised images of 2048 $\times$ 2048 pixels, since no wide field maps from any of the COBRaS observations were in existence. However, the application of the original code upon some of the initial wide field images obtained from the COBRaS Legacy data proved unsuccessful. This was predominantly due to the sheer size of the wide field images which span $\sim$ 16000 $\times$ 16000 pixels. As a result the program was unable to process such a large volume of image pixels and had to be almost entirely re-written and optimised to increase its efficiency. Further additions were made to increase its overall functionality, flexibility and information output. The code was named as the Source Extraction Algorithm for COBRaS (SEAC) and has become the main tool in compiling the source population catalogue for the COBRaS L-band observations. The code is maintained by myself and Danielle Fenech at University College London and is completely open source, made available via GitHub at the following web address: http://github.com/jackmorford/SEAC.

The program has a number of user defined parameters to control different aspects of the code which are edited at the beginning of the script. It performs the following steps:

1. Reads in an astronomical image into a two dimensional NUMPY array, calculates the beam size from the image header and the initial RMS of the entire image using the AIPS task IMEAN.

2. If chosen to do so, a noise map is calculated across the image at a user defined resolution to account for the variation in the noise across the image (this is very important in wide field images where the sensitivity can significantly decrease radially from the image centre, due to primary beam effects).

3. The program cycles through each pixel within the image and every $x, y$ pixel value above a given (seed) threshold level (set by the user to be some factor multiplied by

the RMS) is saved into a PYTHON list.

4. Each of the seed pixels in the list is then run through the floodfill algorithm whereby adjacent pixels above a user defined (flood) threshold level are considered to be part of the same source. In this way each seed pixel is grown into an island which must consist of a number of pixels larger than a given percentage of the beam size for it to be considered[1].

5. Having obtained a given number of islands (each of which represent a detected source), the next part of the code calculates the integrated flux, maximum pixel and weighted position in sky coordinates, the local noise level and their associated uncertainties from one of two possible methods:

   (a) By using the AIPS task JMFIT to fit up to four Gaussian components to each island, from which the various statistics are calculated.

   (b) Or from within the PYTHON environment, summing the pixel flux densities for the island and subtracting it from the local background to obtain the integrated flux density, finding the RMS from around the immediate vicinity of the source and calculating their related uncertainties.

6. In the completion of the island statistics, the script will then utilise the MATPLOTLIB library to plot out each individual source and its associated detected area, as well as the created noise map (if chosen to do so) as well as the entire image with each island clearly labelled as single png files.

7. The png files will be combined into a single pdf file and saved to a user defined directory, as will a table of the derived quantities belonging to each island in text file format.

[1]This was chosen to be 95% of the beam in order to account for low level point like sources within the data.

#### 3.1.1.1 The floodfill algorithm

SEAC utilises the floodfill algorithm in order to identify sources within any given image. It was deemed as the best option in extracting sources from the COBRaS observations due to its low rate of false detections and its flexibility in terms of finding both resolved and unresolved sources. The algorithm takes two user defined inputs; the seed threshold ($\sigma_S$) and the flood threshold ($\sigma_F$). The seed threshold defines the level at which any pixel across the entire image can be first considered as an 'island'. The algorithm then takes each seed pixel on an individual basis and tests the level of the adjacent pixels surrounding it to see if they are above the flood threshold. Every adjacent pixel to the seed pixel above $\sigma_F$ is considered apart of the island. This iterative process is continued until all the adjacent pixels to the island are lower than the flood threshold level. Since we are working with radio maps that have been convolved with a synthesised beam, an island is only generally considered an island if the number of pixels representing it, are larger than the number of pixels that make up the beam area. However, for low level sources, the number of detected pixels with SEAC may be less than the number of pixels covered by the beam size since the true extent of the source may be covered by the noise level of the surrounding pixels.

Logically, the seed threshold must be larger or equal to that of the flood threshold. In AEGEAN, as described by Hancock et al. (2012), the authors set $\sigma_S = 5\sigma$ and $\sigma_F = 4\sigma$, where $\sigma$ represents the local RMS noise around the source. Whilst $5\sigma$ is generally considered as standard practice for a detection threshold, they find that their flood threshold level is sufficient to accurately characterise point-like (and therefore Guassian-like) islands. In the BLOBCAT algorithm however, Hales et al. (2012) set $\sigma_S = 5\sigma$ and $\sigma_F = 2.6\sigma$. They claim that this much lower value of the flood threshold provided a better representation of the true source pixels for faint sources without over flooding into off-source (noise) pixels. Within SEAC, both $\sigma_S$ and $\sigma_F$ are manually set by the user to give an added flexibility to the program.

#### 3.1.1.2 Background noise map

The flood and seed threshold levels are the two most defining parameters affecting the results from SEAC. However, these represented multiples of the background RMS within the image. When dealing with small e-MERLIN L-band images of 512 $\times$ 512 pixels spanning $\sim$ 20 arcseconds in diameter, the variation in the noise level across the image is insignificant. To obtain an accurate noise level from which each source can be detected against, the AIPS task IMEAN is used across the entire image to calculate the mean and RMS brightness. The calculated RMS and mean is used in order to firstly search for the sources within the floodfill algorithm, and secondly in the calculation of the integrated flux density and its associated error. The wide field images produced in the reduction of the COBRaS 21cm data however span $\sim$ 15 arcminutes in diameter, covering the entire field of view of the e-MERLIN array at L-band. In correction for the response of the primary beam across each of the wide field images, the noise level varies significantly from the centre to the outer edges. The variation can be as much as a factor of $\sim$ 3 meaning a single measurement of the noise level across the entire image will not accurately represent the localised RMS of each source.

To compensate for this variation in noise, SEAC can create a separate noise map that mimics the size and shape of the image array, yet its pixel values correspond to the localised noise. The user can input the resolution of the noise map by choosing the number of boxes in the $x$ and $y$ directions from which to calculate the noise level. An example can be seen in Figure 3.1 calculated from an initial wide field image of pointing A from the COBRaS L-band Legacy observations. The original image has been placed onto a 32 $\times$ 32 cell grid with each cell measuring 488 pixels in diameter. This is a small enough resolution that the noise is unlikely to vary significantly between each grid cell, yet large enough such that a reliable estimate of the noise level is calculated even if bright sources are included within the grid cell. It is clear to see the variation in the noise across the wide field image, the shape of which can be attributed to the response of the primary beam of the e-MERLIN array. When cycling through each pixel within the image array, depending on their position in the array, their amplitude is compared to the noise level of the same pixel position in

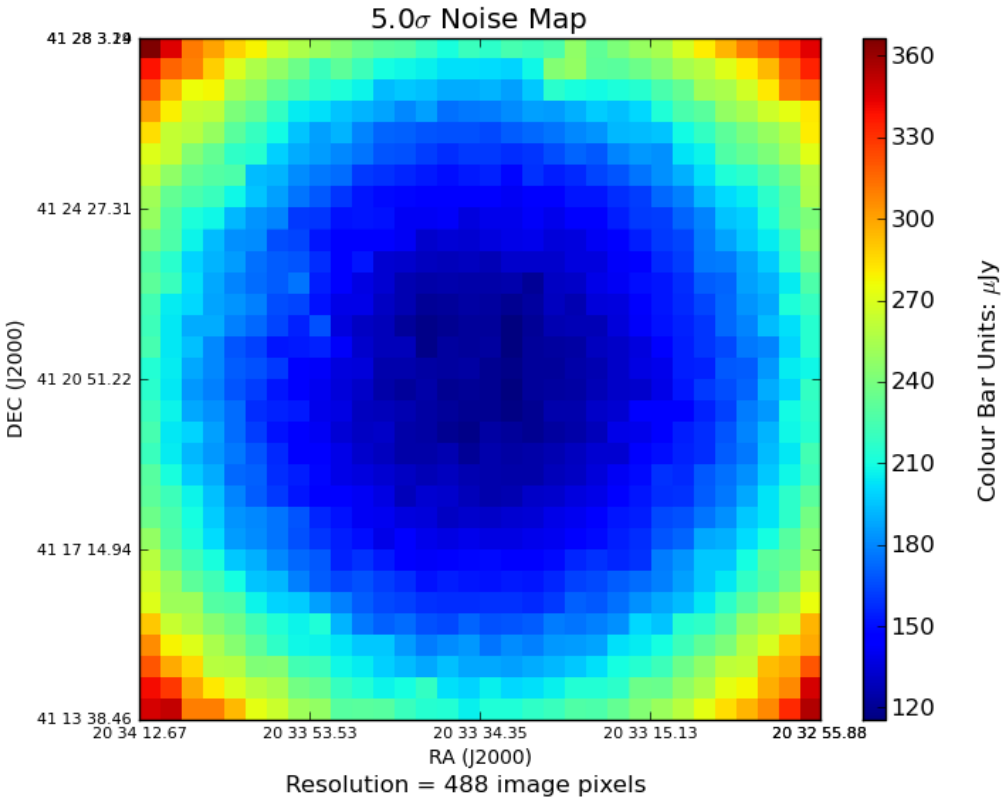



FIGURE 3.1: The calculated noise map created within SEAC in units of 5 × the local RMS. The noise has been calculated in every box of size 488 pixels squared using IMEAN from the input image of CYG-L-A, one of the seven target field pointings taken from the COBRaS L-band Legacy data.

the noise array. Therefore accounting for the variation in noise level across large images meaning each source detected is fairly represented by its local background.

#### 3.1.1.3 Source position determination

Alongside the determination of the integrated source fluxes, the precise position on the sky of each source is of upmost importance. Not only because the position is used to cross-correlate each source with existing catalogues but also because the high resolution obtained from these COBRaS observations has given unrivalled precision as to the true position of each source. Furthermore, the multiple epoch observations may allow for some radial velocity and proper motion measurements of some of Cyg OB2 clusters members.

Within SEAC, there exist several methods to measure the position of each detected source or 'island'. Firstly, when using the AIPS task JMFIT to fit Gaussian models to each island by a least squares method, the task will return the integrated flux of the island and the peak position of the fitted model. If deriving the fluxes and positions from an analysis of the island pixels within PYTHON however, the user can choose to either return the position of the peak pixel position of each island, or the position of the weighted mean pixel. The position of the weighted mean pixel will obtain the most accurate position for resolved, non-gaussian sources, although is more computationally expensive to compute. Using the peak pixel position is also sufficient to accurately represent the position of point-like, unresolved sources.

The weighted mean pixel position $\overline{(x_w, y_w)}$ of a particular island is calculated using the relation:

$$\overline{(x_w, y_w)} = \frac{\sum_{i,j=1}^{N} x_i F(x_i, y_j), y_j F(x_i, y_j)}{\sum_{i,j=1}^{N} F(x_i, y_j)} \tag{3.1}$$

where $F(x_i, y_j)$ represents the flux of the pixel $(x_i, y_j)$, and $N$ represents the total number of pixels belonging to the island. Similarly, the error on this weighted mean pixel position

is given by:

$$\delta\overline{(x_w, y_w)} = \sqrt{\sum_{i,j=1}^{N} \left( \frac{x_i, y_j}{F(x_i, y_j)} \right)^2} \tag{3.2}$$

Each of these methods will only return the $x$ and $y$ pixel position of each source. SEAC then converts this pixel position into the true position on the sky in astronomical coordinates. This is accomplished by obtaining the sky coordinates (in degrees) according to the central pixel position in the image, and the pixel increment in both the $x$ and $y$ direction from the AIPS image header. By calculating the distance offset in radians between the peak pixel position and the central pixel position in both the x and y direction, the Right Ascension (RA) and Declination (DEC) of the source can be calculated in radians which can be further converted into units of degrees or hours, minutes, seconds (hms).

#### 3.1.1.4 Source flux determination

The source integrated flux density and its associated error can either be derived using the AIPS task JMFIT or via an analysis of the detected island pixels. The flux calculated with JMFIT can vary significantly depending on the input parameters. The AIPS task benefits hugely from some prior information regarding the general source shape and is therefore tailored towards point-like, un-resolved sources. In regards to resolved (un-Gaussian like) sources, JMFIT struggles to accurately represent the source and generally results in an un-reliable integrated flux density. Great care must be taken if using JMFIT to determine the source fluxes of resolved or extended sources. Thus, in the case of resolved sources, a better option is to use an analysis of the determined island pixels. The integrated flux of a given island is calculated as follows:

$$F = \frac{1}{N_{beam}} \left( \sum_{i,j=1}^{N} F(x_i, y_j) \right) - \frac{\overline{x}_{LB} N}{N_{beam}} \tag{3.3}$$

where $\overline{x}_{LB}$ is the mean of local background, $N_{beam}$ is the number of pixels representing the size of the beam and $N$ is the number of pixels belonging to the island. The local

background should represent the general noise level in the vicinity of the source. Within SEAC, this can be calculated individually for each source, or over the entire image. Calculating the mean and RMS of the local background around the immediate vicinity of each source will give a much more accurate representation of the true background for each source. SEAC includes an option to change the size of this local background surrounding each island but as a default uses a box of $4 \times 4$ arcseconds centred around each island. Note that the mean and RMS noise calculated from the local background does not include the pixels that make up the island. The associated error in the integrated flux calculation is given by:

$$\delta F = RMS_{LB} \times \frac{N}{N_{beam}} \tag{3.4}$$

where the $\mathrm{RMS}_{LB}$ is the root mean square noise calculated from the local background. The reliability of the integrated flux of a given source, calculated in Equation 3.3 depends purely on the reliability of the floodfill algorithm. The detected pixels that contribute to the integrated flux must be a fair representation of the true source on the sky and therefore the seed and flood threshold levels that dictate the detected area must be chosen with care.

#### 3.1.1.5 SEAC outputs

Figure 3.2 shows a typical example of the output plots from SEAC. The program will merge each of the created png files into a single pdf file for convenience. In this particular example, the input image of $2048 \times 2048$ pixels covered a region of the sky sampled by pointing E of the COBRaS L-band Legacy observations. Within the image there exist two astronomical sources that SEAC has identified and then calculated their mean weighted positions on the sky as well as their integrated fluxes. The four bottom panels of Figure 3.2 give an example of the two detected sources with SEAC. The left panel, in either case, represents the detected pixel area determined by the floodfill algorithm and the right panel shows the actual image intensities (in the $z$ direction). Further information regarding the source integrated flux, the image beam size, the source peak pixel flux and the sky position are shown in each of these 'snapshot' images. A further plot (show in the top left panel

of Figure 3.2) shows the entire input image with the positions of the two detected islands clearly labelled.

SEAC further outputs a text file containing the derived information of each of the detected islands. This includes the position of the source on the sky, its integrated flux, the RMS noise within the local background, the signal-to-noise of each source and their associated uncertainties. The text file is formatted in such a way that a further script used to cross-correlate the information with pre-existing catalogues (see Section 3.2) can easily read the necessary information from which to perform the cross-correlation.

### 3.1.2 Search and destroy - SAD

In order to both validate and compare the results obtained using SEAC, the inbuilt AIPS task Search and Destroy (SAD) can also be implemented upon the COBRaS L-band data. SAD attempts to find sources whose peak pixel flux is brighter than a given level. Those identified that are contiguous become merged into one island from which JMFIT is used to perform a least square Gaussian fit to the identified pixels. The derived parameters including the position and integrated flux of each island are calculated directly from the Gaussian model which are then written into a Model Fit (MF) table, attached to the image. JMFIT can fit up to a maximum of four Gaussian components to each island but returns reliable positions and flux densities for point-like, Gaussian sources with a good signal-to-noise ratio.

For a direct comparison between the two source extraction methods, each algorithm was implemented on four sources found within the COBRaS L-band field of view. The sources were chosen in order to give a fair representation of the different types of radio sources within the COBRaS data set. This included a variation in the signal-to-noise ratio, peak-pixel flux, and the size of each source on the sky with respect to the beam size. Figure 3.3 shows two plots of each of the four chosen sources. Firstly a plot created using the AIPS task KNTR, from which the beam size (found in the bottom left hand corner of each image) can be directly compared to the extent of the source on the sky. Secondly, on the right hand column of Figure 3.3, the output plots from SEAC are shown, including

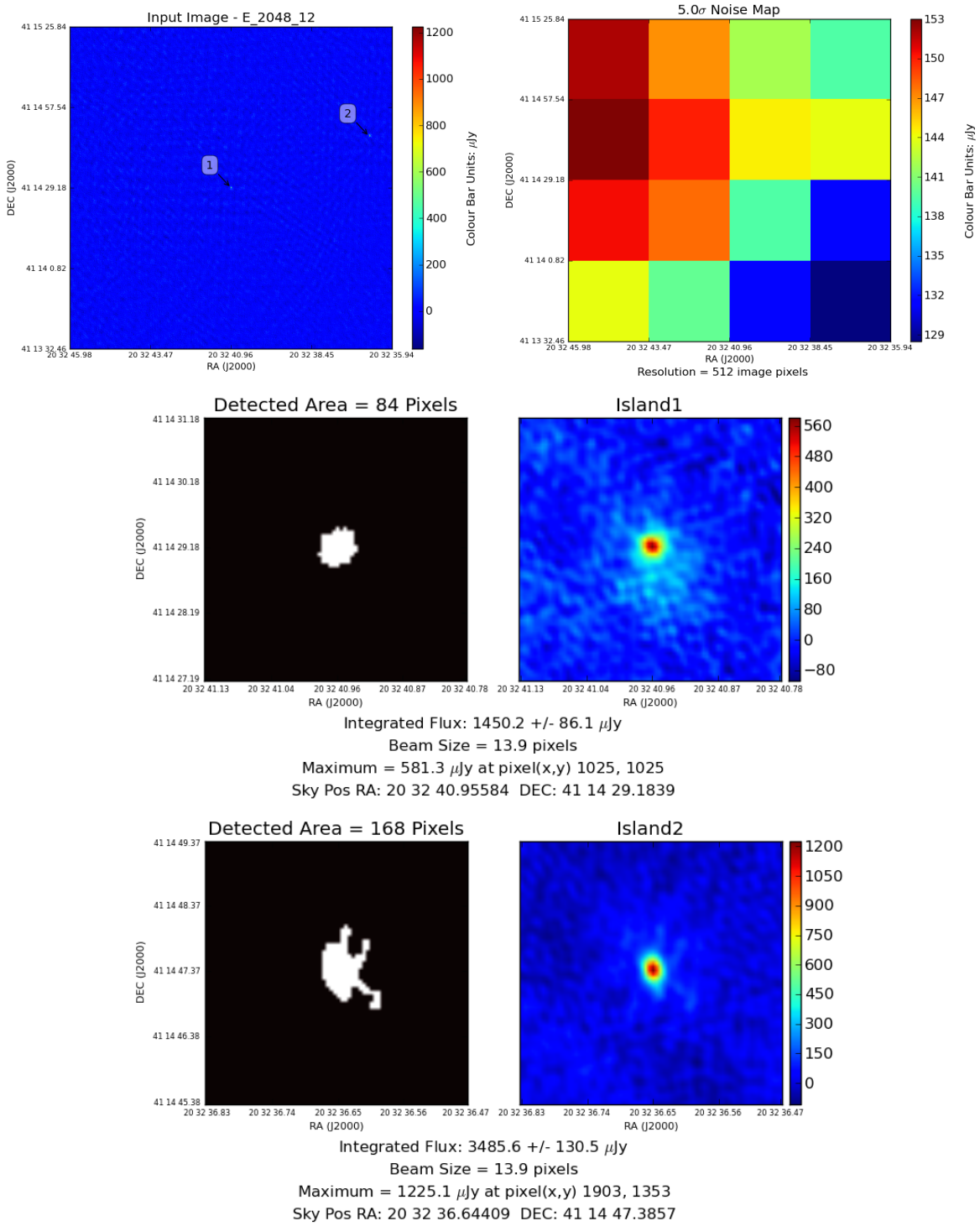



FIGURE 3.2: A typical example of the output plots from SEAC. The program was run with a seed and flood threshold of $5\sigma$ and $4\sigma$ respectively on a 2048 $\times$ 2048 image of a region of the sky within the field of view of pointing E from the COBRaS L-band observations. The two sources within the image which were detected are shown in the bottom four panels. The top left panel shows a plot of the input image with the two detected sources highlighted appropriately. Additionally, the top right image shows the calculated noise map at a resolution of 512 pixels.

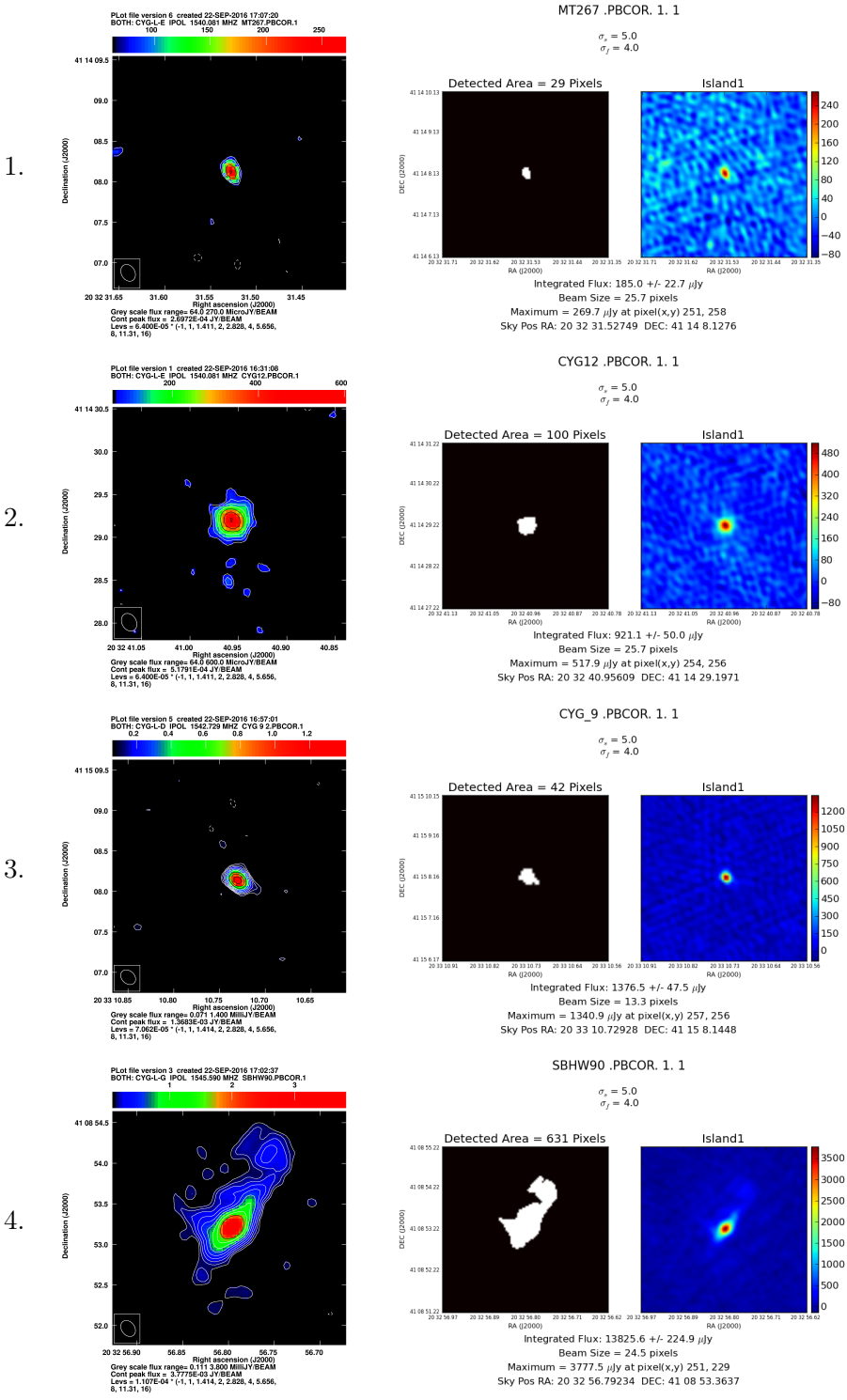



FIGURE 3.3: Four different sources within the COBRaS L-band field of view, the left hand column shows the plots made with the AIPS task KNTR whilst the right side shows the output png files from SEAC. 1. Cyg OB2 #12; 2. A11 (MT267); 3. SBHW90; 4. Cyg OB2 #9. The positions and integrated fluxes for each of the four sources, as determined by SEAC and SAD, can be seen in Table 3.1.

TABLE 3.1: To highlight the differences between the positions and integrated fluxes as obtained via SEAC and SAD for four different sources of varying brightness and size. Both algorithms are detecting sources down to 5 × the image RMS (5$\sigma$; SEAC is using an additional flood threshold $\sigma_F$=4$\sigma$). The corresponding images of these four sources can be seen in Figure 3.3.

| Source Name | 1. A11 | 2. CYG#12 | 3. CYG#9 | 4. SBHW90 |
|---|---|---|---|---|
| RA (J2000)* | 20 32 31.527 | 20 32 40.958 | 20 33 10.729 | 20 32 56.792 |
| DEC (J2000)* | 41 14 08.128 | 41 14 29.220 | 41 15 08.145 | 41 08 53.364 |
| Resolved? | No | Yes | No | Yes |
| S/N* | 9 | 36 | 51 | 312 |
| Position Offset | 37 mas | 6 mas | 11 mas | 143 mas |
| $F_{SEAC}(\mu Jy)$ | 185 ± 22 | 921 ± 50 | 1377 ± 48 | 13826 ± 225 |
| $F_{SAD}(\mu Jy)$ | 280 ± 39 | 1131 ± 72 | 1424 ± 142 | 8659 ± 108 |

*As measured using the results from SEAC.

the detected area of each island obtained using the floodfill algorithm. Having run both SEAC and SAD upon each of the four images, Table 3.1 shows the difference between their respective results including the offset in the derived source position from each algorithm and the integrated flux density of the source found by either method.

The image of source number 3, the colliding wind massive star binary known as Cyg OB2 #9 (see Chapter 5 for more details), is compact, includes a good signal-to-noise ratio and is only marginally resolved within the e-MERLIN L-band observations. As a result, this source can be well represented by a Gaussian model leading to a reliable flux density and position as determined by SAD. With reference to Table 3.1 the derived flux densities from both SAD and SEAC are within $\sim$ 3 % of one another showing the reliability of either source extraction method for this type of source. The similarity between the derived flux densities between the two extraction algorithms starts to break down when different types of sources are encountered. The blue hypergiant (and candidate Luminous Blue Variable) star Cyg OB2 #12 (see Chapter 4 for further details) shown as source number 2 in both Table 3.1 and Figure 3.3 still has a decent signal-to-noise ratio yet is mostly resolved within the COBRaS L-band observations. Whilst it can be still represented by a Gaussian model, the source structure leads to a slightly larger ($\sim$ 20%) difference between the derived flux densities of SEAC and SAD.

A11 (source number 1), a binary system with an O7.5 III primary, represents a very compact, un-resolved source of a low signal-to-noise ratio. The difference in the derived flux density from both SEAC and SAD is a factor of $\sim 1.5$, a significant difference for a source with a low signal-to-noise. If the floodfill threshold, $\sigma_f$, is dropped to 3.0 $\times$ the image RMS, SEAC provides an integrated flux density of $219 \pm 26$ $\mu$Jy for A11. This shows that for islands with a low signal-to-noise ratio, the full extent of the source can be lost within the surrounding noise level and SEAC can potentially underestimate their flux densities. In contrast, SAD is unlikely to underestimate the flux density of low signal-to-noise source, however with such little information regarding the true extent of the source on the sky, the Gaussian fit is unlikely to well represent the source and may potentially lead to un-reliable flux density measurements.

At the other end in the spectrum of source types, SBHW90, an unclassified object represents a resolved, bright radio source with lots of extended structure. The flux density derived from SEAC is $\sim 1.6$ $\times$ larger than that derived from SAD. This highlights the difficulty in accurately representing resolved, extended sources with up to four Gaussian components, as in the case with SAD. This small study into the difference between the two source detection algorithms shows that SEAC will be better implemented upon the COBRaS 21cm radio maps. SEAC has been shown to be versatile, delivering reliable source positions and flux densities for a range of source types that differ in their structure, signal-to-noise, and extent on the sky.

Further comparisons between the source detection efficiency and flux determinations obtained between JMFIT (i.e. SAD) and pixel-by-pixel (PP; i.e. SEAC) methods were made in Peck (2014). By testing these methods on simulated point-like and resolved Gaussian sources at a range of signal-to-noise ratios, Peck (2014) found that the PP method was shown to perform better in terms of flux extraction than JMFIT for sources with a SNR $\geq 5$ (despite the non-Gaussian nature of the PP method). Moreover, for sources of a low SNR, the PP method was found to suffer in measuring their flux densities due to the intrinsic property of source declaration as a function of the surrounding noise. JMFIT on the other hand was found to consistently underestimate the flux by $\sim 4\%$ at higher SNRs whilst also producing a number of false detections for SNR $\leq 5$.

## 3.2 Catalogues and cross-correlation techniques

The purpose built Source Extraction Algorithm for COBRaS (SEAC) returns a text file containing the derived information of each detected source within the image on which it was implemented. Along with the flux density, signal-to-noise noise, RMS of the local background and their associated uncertainties, the text file contains the positional information of each detected source in units of both degrees and hours, minutes, seconds (hms). This position is used to compare directly to the positions of previously identified sources that lie within the Cygnus OB2 association. Each entry of the COBRaS L-band All Source Catalogue (CLASC) matched in both Right Ascension (RA) and Declination (DEC) to a previous catalogue can be automatically identified and depending on the pre-existing information of the particular source, it may also be classified. In order to maximise the number of correct matches and therefore identifications the catalogue from which the CLASC is cross-correlated with must be as exhaustive as possible, including all previously detected sources that lie within the field of view of the COBRaS L-band observations. The following section describes the steps taken in amalgamating the Cygnus OB2 super catalogue, forming the base catalogue from which the CLASC and indeed all future COBRaS catalogues (i.e. C-band) can be cross-correlated with. Other catalogues which provide further cross-correlation with CLASC are described including the formulation of two candidate OB star catalogues.

### 3.2.1 The Cyg OB2 super catalogue

The amalgamation of the Cyg OB2 Super Catalogue was performed in the Tools for OPerations on Catalgoues and Tables (TOPCAT; Taylor, 2005) environment by utilising scripting procedures that make use of the Starlink Tables Infrastructure Library Tool Set (STILTS; Taylor, 2006). Each catalogue, including source information within the Cyg OB2 region, was found using the VizieR Service which could be downloaded in a Virtual Observatory (VO) table format, ready for manipulation within TOPCAT. A total of 32 catalogues were found to contain information on sources within the same region of the sky as occupied by the Cyg OB2 association. These are listed in Table 3.2, along with the number of sources

they contain. Catalogues 1 to 24 were originally combined by Peck (2014) to form the Cyg OB2 Super catalogue. However since then, a further 6 catalogues (25 to 30 in Table 3.2) have been added, increasing the number of individual sources from 6500 to 14355. Catalogue numbers 31 and 32 include the 2MASS sources within a 24 arcminute radius of the centre of the association and the photometric observations of Cyg OB2 as conducted by Guarcello et al. (2012), respectively. Since they each contain such a large quantity of sources, they have been made absent from the Cyg OB2 Super Catalogue and will be treated individually in the cross-correlation process with the CLASC.

A JYTHON script was used to perform the concatenation of the 30 individual catalogues. JYTHON allows for the PYTHON language to be implemented in a JAVA platform, such as the STILTS package. Before concatenating the catalogues and cross-matching sources by their sky coordinates, the common identifiers of MT91 (Massey and Thompson, 1991) and AFM2007 (Colombo et al., 2007b) were first used to cross-match any duplicate sources between the catalogues. A total of 10 catalogues included the MT91 identifier which were first cross-correlated and concatenated into a single MT91 catalogue consisting of 1047 sources. Three catalogues included the AFM2007 identifier, which were subsequently concatenated into a single AFM2007 catalogue containing 1723 sources. The remaining catalogues were combined together and their sources cross-matched with one another by using their RA and declination (J2000) in degrees. In order to account for the difference in the accuracy of the positions of each source from different catalogues, the sky coordinates were matched to within a radius of 3 arcseconds of one another. This was accomplished using the STILTS command 'tmatch2'.

The complete concatenation and cross-matching of catalogues 1 to 30 (as listed in Table 3.2) formed the Cyg OB2 Super Catalogue, containing 14546 sources. Duplicates still existed since some sources with multiple components were not denoted consistently from catalogue to catalogue. The JYTHON script further removed these duplicates by cross matching each source within the entire catalogue a second time, whilst retaining the column information for either duplicate row. This shortened the output catalogue to 14413 sources. To further remove any sources lying outside the region covered by the Cyg OB2 association, a cone search was applied to the catalogue. This filtered out all sources, whose sky coordinates

TABLE 3.2: A list of the previous catalogues found using the VizieR Service, all of which include sources from the region covered by the Cyg OB2 association.

| # | **Reference** | **Waveband** | **# of Sources** |
|---|---|---|---|
| 1 | Massey and Thompson (1991) | Visible | 801 |
| 2 | Condon et al. (1998) | Radio | 72 |
| 3 | Pigulski and Kołaczkowski (1998) | VIR | 288 |
| 4 | Comerón and Torra (2001) | JHK | 320 |
| 5 | Comerón et al. (2002) | KH | 85 |
| 6 | Setia Gunawan et al. (2003) | Radio | 239 |
| 7 | Wolff et al. (2007) | Visible | 13 |
| 8 | Colombo et al. (2007b) | X-Ray | 1003 |
| 9 | Colombo et al. (2007c) | X-Ray | 147 |
| 10 | Martí et al. (2007) | Radio | 153 |
| 11 | Kiminki et al. (2007) | Visible | 303 |
| 12 | Kiminki, McSwain, and Kobulnicky (2008) | Visible | 17 |
| 13 | Vink et al. (2008) | Visible | 71 |
| 14 | Kiminki et al. (2009) | Visible | 22 |
| 15 | Wright and Drake (2009) | X-Ray | 1696 |
| 16 | Skiff (2009) | Various | 150 |
| 17 | Kobulnicky, Gilbert, and Kiminki (2010) | Visible | 17 |
| 18 | Kiminki et al. (2012) | Visible | 21 |
| 19 | Kiminki and Kobulnicky (2012) | Visible | 46 |
| 20 | Kobulnicky et al. (2012) | Visible | 28 |
| 21 | Comerón and Pasquali (2012) | Visible | 240 |
| 22 | Radio Master Catalogue* | Radio | 2850 |
| 23 | Galactic O Star Catalogue** | Various | 15 |
| 24 | Simbad Database*** | Various | 2077 |
| 25 | Rauw (2011) | X-Ray | 453 |
| 26 | Guarcello et al. (2013) | Various | 1843 |
| 27 | Kobulnicky et al. (2014) | Visible | 50 |
| 28 | Wright et al. (2014) | X-Ray | 7924 |
| 29 | Wright, Drew, and Mohr-Smith (2015) | Various | 167 |
| 30 | Wright et al. (2016) | Various | 873 |
| 31 | Cutri et al. (2003) | JKH (2MASS) | 27090 |
| 32 | Guarcello et al. (2012) | RIZ | 64157 |

*Radio Master catalogue is a combination of two previous catalgoues from Dixon (1970) and Kühr et al. (1979).
**Galactic O star Catalogue was taken from Maíz-Apellániz et al. (2004).
***Simbad Database catalogue was constructed using a VO search results with a radius of 24 arcminutes, centred on the Cyg OB2 association.

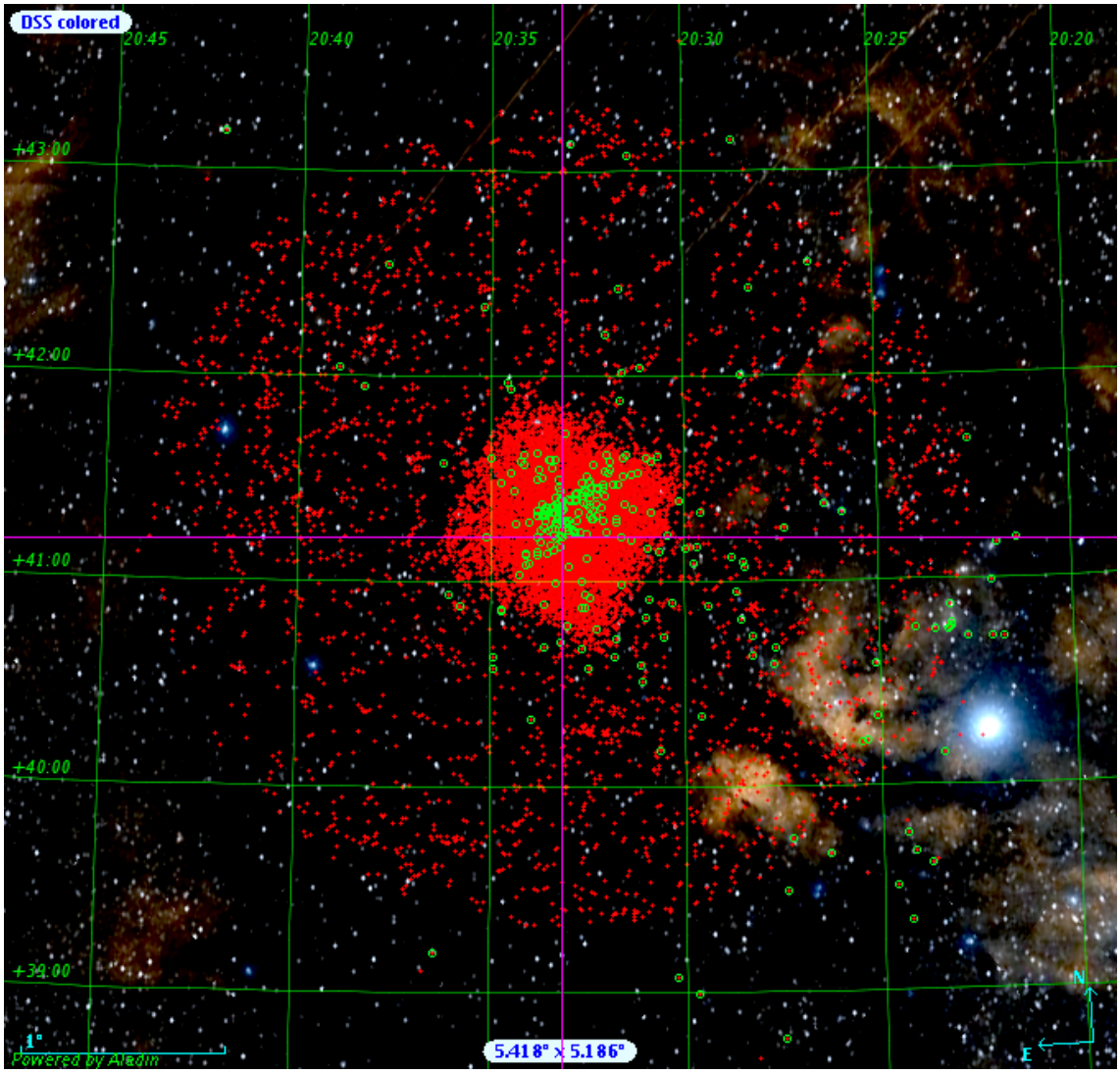



FIGURE 3.4: The positions of the 14355 sources within the Cyg OB2 Super Catalogue (red crosses), overlaid onto a colour Digitized Sky Survey (DSS) image centred on J2000 coordinates RA: $20^h33^m10.8^s, DEC : 41°13'12.0''$. The green circles represent the known OB type stars within the catalgoue. The image spans an area on the sky of $5.418° \times 5.186°$ and was created using the Aladin Sky Atlas.

were not within a 3° radius of the centre of the Cyg OB2 association (in J2000 coordinates - RA: $20^h33^m10.8^s$, DEC: $41°13'12.0''$) and gave rise to the final Cyg OB2 Super Catalogue, containing 14355 sources. Figure 3.4 shows the positions of each of the 14355 sources in the catalogue, overlaid onto a Digitized Sky Survey (DSS) image.

#### 3.2.1.1 Cyg OB2 OB star catalogue

The Cyg OB2 Super Catalogue contains a large variety of sources across a number of different wavebands (see Table 3.2). The key science goals of the COBRaS project involve an analysis of the radio emission from the massive stellar members of the association, i.e. the OB-type stars. Thus, a separate catalogue, containing only information upon the OB members of the association is required. Again, STILTS commands within a JYTHON script can be utilised to cycle through the columns of the super catalogue, identifying all of those with spectral type information within them. This information is copied into a single column that contains the spectral type information for all rows for which it exists. By then selecting the rows with spectral types of either type 'O' or 'B', the Cyg OB2 OB star catalogue is created. A total of 300 OB stars are found within the super catalogue and their sky positions are indicated by the green circles within Figure 3.4.

#### 3.2.1.2 Cyg OB2 OB candidate catalogue

Further work by Peck (2014) was done to derive an OB star candidate catalogue from the rows within the super catalogue that contained columns with JHK magnitude values. In plotting together the (J-H) and (H-K) values, for a large number of sources from the 2MASS catalogue and acquiring follow up spectroscopic observations, Comerón et al. (2002) found that early type stars dominate a particular region of the (J-H, H-K) diagram, that they labelled as the 'blue group'. This 'blue group' of early-type stars was limited to those with (H-K) values $< 0.5$. This technique was applied to the original Cyg OB2 super catalogue by Peck (2014) which resulted in an extra 317 previously unclassified, OB candidate stars.

In updating the Cyg OB2 super catalogue to include a further 6 catalogues containing 7681 sources, the technique of selecting OB candidate stars from their position on a (J-H,

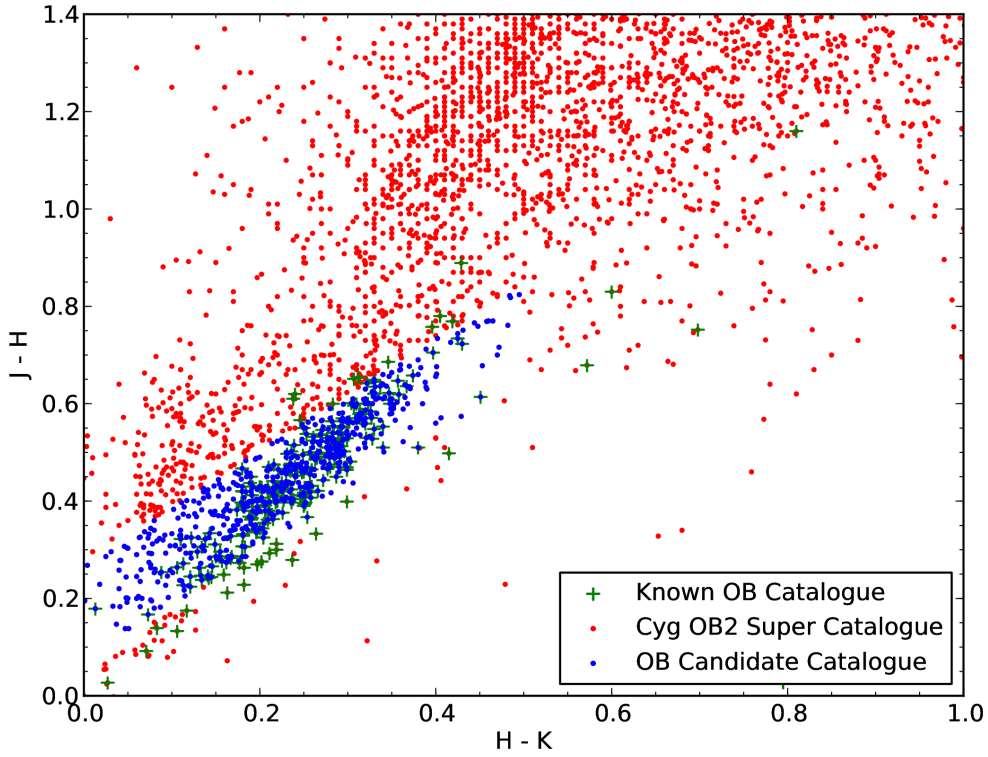



FIGURE 3.5: (J-H, H-K) diagram of the sources from the Cyg OB2 super catalogue that contain JHK magnitude information. There includes 4252 sources possessing JHK magnitude values, a linear progression is fitted to the green crosses highlighting the known OB stars within the region. The blue dots are declared to be OB candidate stars, since these have (H-K) values $< 0.5$ and fall within a loci of 0.1 magnitudes around the linear fit to the green crosses.

H-K) diagram was once again utilised. In first cycling through each of the columns within the super catalogue to find any J, H and K band magnitudes, new columns were created to contain the (J-H) and (H-K) values for each row. By fitting a linear function to the known OB stars within the super catalogue on the (J-H, H-K) diagram, the rest of the sources found within a loci of 0.1 magnitudes of this line (and under the condition that the (H-K) value is $< 0.5$) were considered to be OB candidate stars. Figure 3.5 shows the OB candidate stars (shown as blue dots) as selected from (red) sources from the super catalogue. A total of 343 OB candidate sources have been identified, adding to the 300 previously known OB-type stars within the Cyg OB2 association (which are highlighted by green crosses in Figure 3.5).

### 3.2.2 Other catalogues considered for cross-correlation

Catalogues 31 and 32 (as listed in Table 3.2) each contain a vast number of sources, significantly more than the total number found in the Cyg OB2 super catalogue. For reasons of computational efficiency, these two catalogues are therefore treated separately within the cross correlation process and are briefly described below.

#### 3.2.2.1 2MASS catalogue

The Two Micron All Sky Survey (2MASS; Cutri et al. 2003), performed a whole sky survey in three near-infrared wavebands (JHK). The 2MASS point source catalogue contains accurate positions and fluxes of $\sim$ 300 million stars and other unresolved objects, whilst the extended sources catalogue contains positions and magnitudes for over $10^6$ galaxies and other nebulae. In tracing a circle with a 24 arcminute radius and centred upon the Cyg OB2 association, the 2MASS point source catalogue returns 27090 sources. A smaller radius (as opposed to the 3° radius used for source in the Cyg OB2 catalogue) was required due to the sheer number of sources found within the 2MASS catalogue, many of which falling outside the area of the sky covered by the COBRaS observations. This is over a factor of two larger than the amalgamated Cyg OB2 super catalogue and will provide an excellent reference from which to cross-correlate with the CLASC. Furthermore, the

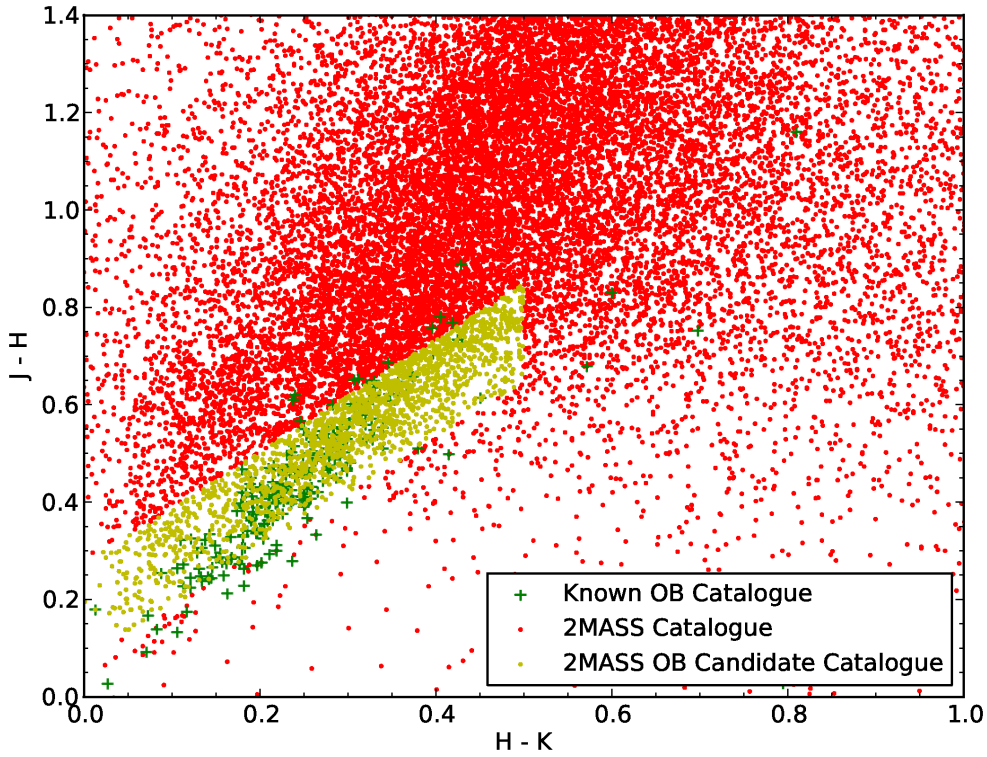



FIGURE 3.6: (J-H, H-K) diagram of the sources from the 2MASS catalogue that lie within a 24 arcminute radius of the centre of the Cyg OB2 association. There include 27090 sources, the yellow dots highlight those declared to be OB candidate stars whilst the green crosses highlight the known OB stars from the Cyg OB2 super catalogue.

2MASS survey contains JHK magnitudes for each of these 27090 sources within the region of the sky covered by the Cyg OB2 association. In a similar fashion to the super catalogue, a 2MASS OB candidate catalogue can be derived by identifying those sources belonging to the 'blue group' in the (J-H, H-K) diagram. Figure 3.6 shows the (J-H, H-K) diagram for the 2MASS catalogue, with the 1930 green dots, highlighting the OB candidate sources derived using the same technique as that described previously.

#### 3.2.2.2 Optical photometry GTC/OSIRIS catalogue

A final catalogue also to be used within the cross-correlation process with the CLASC is the large optical catalogue from Guarcello et al. (2012) (hereafter G2012). Their photometric observations in the three optical *r, i, z* filters taken with the Optical System for Imaging and low Resolution Integrated Spectroscopy (OSIRIS) mounted upon the 10.4m Gran Telescopio CANARIAS (GTC) instrument provides the largest optical catalogue of the Cyg OB2 association. The catalogue consists of 64157 sources down to a mass of M = 0.015 $M_{\odot}$ and covers a $41' \times 41'$ region, centred on the middle of the Cyg OB2 association. Containing such a vast number of sources, this catalogue will be of great value when cross-correlating the CLASC. Figure 3.7 shows the positions of each of the 64157 sources (shown in blue) plotted against a background DSS image with sources from the Cyg OB2 super catalogue shown in red.

### 3.2.3 Cross correlation script

The cross-correlation of the CLASC with the various catalogues of previous source detections within the Cyg OB2 region is achieved through the use of a JYTHON script. Various STILTS commands are used to access the positional information within the catalogue (stored as a VO table) that the CLASC is to be cross-correlated against. The output from SEAC (and therefore the information within the CLASC) is stored as a text file. PYTHON commands within the JYTHON cross-correlation script find the positional information of each detected source from the COBRaS L-band observations and match the sky coordinates against those within the catalogue of previously detected sources. The script requires a

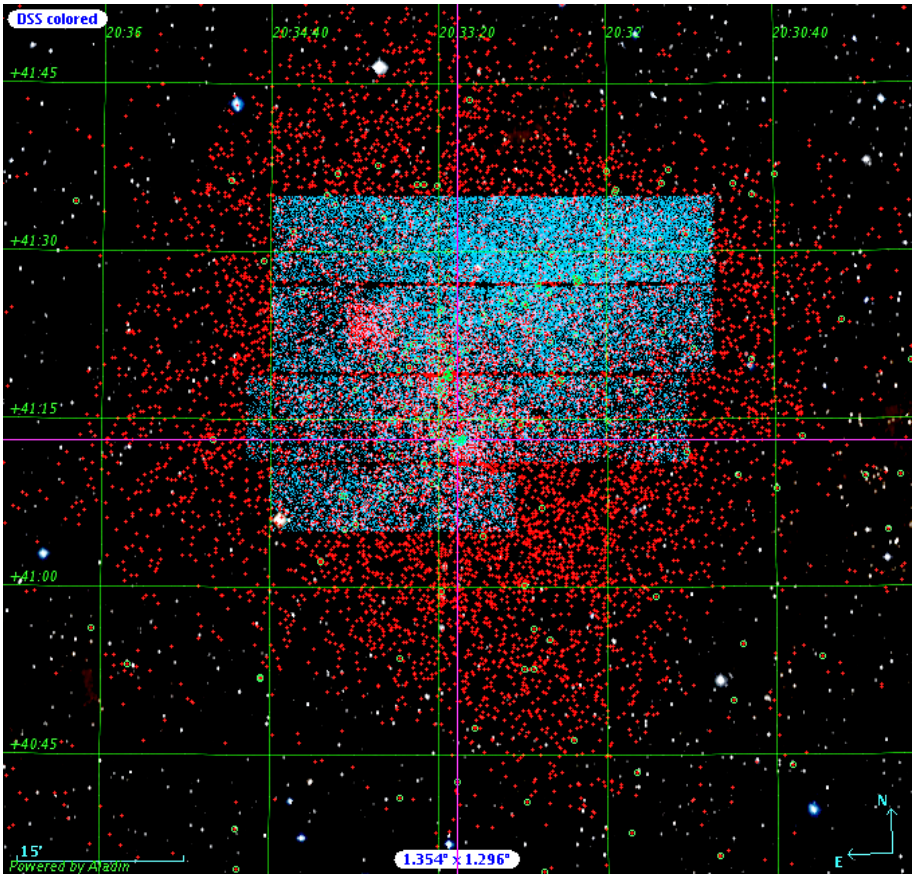



FIGURE 3.7: The positions of the 64157 sources from the GTC/OSIRIS observations from Guarcello et al. (2012) (blue dots), overlaid onto a colour Digitized Sky Survey (DSS) image centred on J2000 coordinates RA: $20^h33^m10.8^s$, DEC: $41°13'12.0''$. The green circles represent the known OB type stars from within the Cyg OB2 super catalogue, whose other source members are shown in red. The image spans an area on the sky of $1.354° \times 1.296°$ and was created using the Aladin Sky Atlas.

positional error given in arcseconds in order to account for the uncertainty in the sky coordinates associated with the sources within either catalogue.

The positional information associated with the COBRaS L-band sources will be accurate to $\sim$ 40 mas (based upon the resolution of the wide field images). However the accuracy of the source positions in a given catalogue from which the cross-correlation will be performed against will depend on the pointing accuracy of the telescope used to obtain those observations. At least initially, a conservative approach will be taken, by cross matching each of the sources within CLASC to a positional accuracy of 3" of the sources found within the Cyg OB2 super catalogue, the 2MASS catalogue and G2012 catalogue. The cross-correlation script will calculate the offset in arcseconds between a given match, after which the likelihood of a given match can be made from an assessment the positional accuracy of the source (from a previous detection with a given instrument) and the extent of the offset between the two positions. In order to display the information gathered from the cross-correlation routine, the script saves this information in a tabulated format to a text file.

## 3.3 COBRaS L-band source population catalogue

This section aims to discuss the methodology and techniques used in deriving the final COBRaS L-band All Source Catalogue. By first describing the steps taken in detecting sources within the seven wide field images and eliminating any potential false detections, the positions and integrated fluxes as measured across both observation epochs is presented in Table 3.3. By then cross-correlating the positions of each of the sources against those found in the catalogues described in Section 3.2, each object will be categorised and classified where possible. Each identified object will be discussed individually.

### 3.3.1 Source detection

The wide field images of each of the seven L-band pointings (see Section 2.4.3 for further details on their production) comprised of 19784 $\times$ 19784 pixels (with the exception of

pointing G), where each pixel corresponds to an area of 0.04" × 0.04" on the sky. When imaged the resulting wide field maps covered an area on the sky of $\sim 0.22°$ in RA and $\sim 0.22°$ in declination. The distance between each of the pointing centres was smaller than the areas covered by each of the pointing maps, as a result some areas of the sky were covered by two or even three of the wide field images.

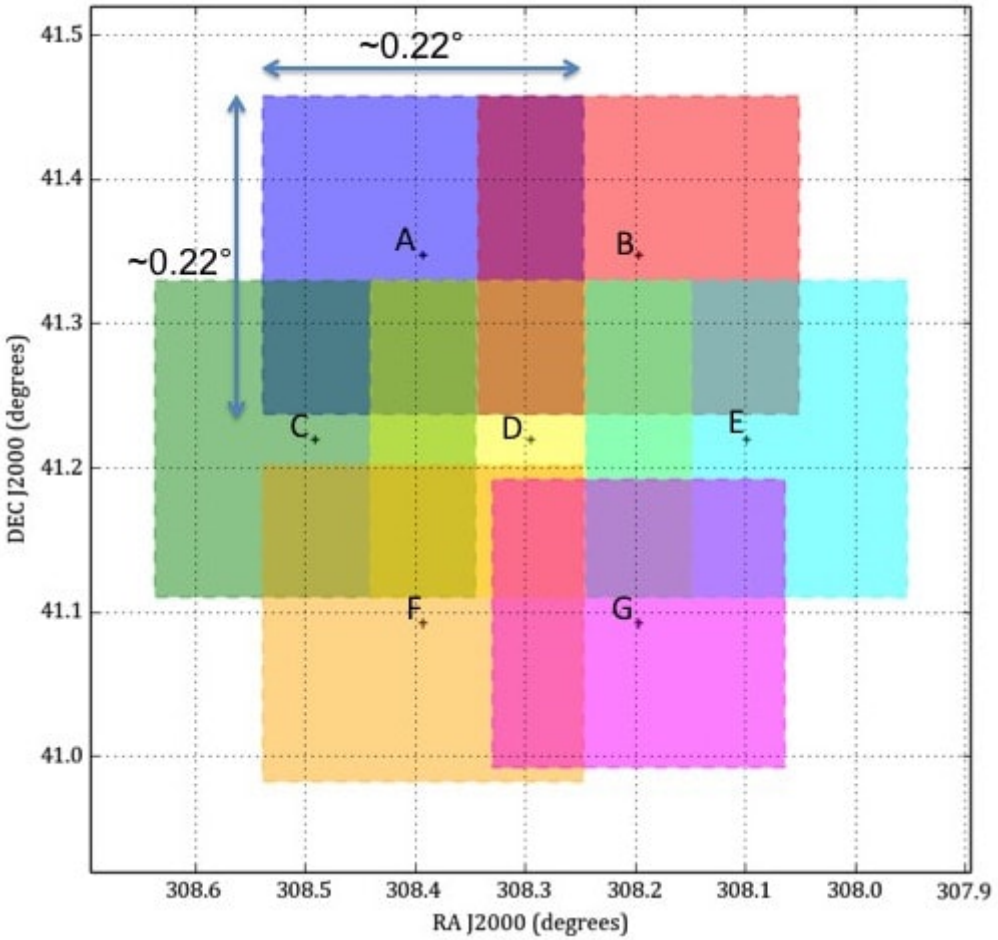



FIGURE 3.8: To highlight the regions covered by the seven wide field COBRaS L-band images on the sky. Each pointing centre is highlighted by a cross and labelled alphabetically from A though to G. Note the slightly smaller size of the wide field image representing pointing G (see Section 2.4.3 for further details).

The positions of the seven pointing centres and the area on the sky covered by the wide field images can be seen in Figure 3.8. The fact that the wide field images overlap means that some of the same sources will appear in multiple pointings (a maximum of three). This is beneficial to the source extraction process since any (potential) 'false' detections

by SEAC in a given pointing can be easily identified if its position also lies within another pointing. If the source is not detected in the same position on the sky within the other pointing, then it can be considered as a false detection, likely due to noise or artefacts associated with the image (assuming its signal-to-noise ratio isn't significantly larger than $5\sigma$).

The following describes the process used to derive the CLASC from the use of SEAC upon each of the seven individual wide field images representing the different pointings associated with the COBRaS L-band observations.

1. SEAC was run on each of the wide field images with a seed threshold, $\sigma_s = 5.0$ and a floodfill threshold, $\sigma_f = 3.0$. A noise map was produced at a resolution of 520 pixels, ensuring potential islands were tested against the RMS calculated from their local pixels.

2. The information from the resulting text file outputs from SEAC (for each of the seven pointings) were combined into one master table, resulting in a total number of 119 islands across the seven wide field images.

3. By inspecting the output images (from SEAC) of the 119 islands, 7 were found to be of a low SNR ($\lesssim 3.0$), within the immediate vicinity of a bright source and clearly associated with surrounding artefacts. As a result these were identified as false detections, reducing the number of islands to 112.

4. Note that in four cases, two (or more) islands were actually found to be a part of the same source. An example is the system of Cyg OB2 #5, which is known to contain two radio components (see Chapter 5 for further details). In each case, the components associated with a given source are labelled alphabetically in order of RA (e.g. 1a and 1b).

5. The positional information of the 109 sources were then cross matched with one another to identify duplicates (i.e. the same source appearing in two or more pointings). 23 duplicates were identified, bringing the total number of individual detections down to 86.

6. A total of 47 of these 86 sources were detected at a signal-to-noise ratio, SNR $<$ 5.0 and the majority of these were found towards the outer edges of their pointing centres.

7. The positional information of a given source, detected within a particular pointing, was considered in relation to the other pointing centres, i.e. the distance between the source position and its own pointing centre was compared to the distance between the source position and that of another pointing centre. If a source was found to lie closer to the pointing centre that wasn't that in which it was first identified, then it should be clearly detected in the other pointing. Therefore for all sources with a S/N $\lesssim$ 5.0, which should have been detected in a separate pointing (since they were positioned closer to its centre), were considered to be false detections and were discarded. This eliminated a further 25 sources.

8. 61 sources then remained, 24 of which had a measured SNR $<$ 5.0. Moreover, the majority of these 24 sources were positioned within the outer regions of their respective pointings and still could not yet be classified as definite detections.

9. The RMS noise level at a given position in the wide field images could be up to a factor of $\sim$ 1.5 $\times$ larger (in the outer edges of the field; see Section 2.4.3 for further details) than the RMS noise level at the same position within a 512 $\times$ 512 pixel image (at the same resolution). Considering this, each of the 24 low SNR sources detected within the wide field images with SEAC, were imaged individually. This was to check if they were real detections, or simply a function of the added noise (and artefacts) inherent to the wide field images. This process eliminated a total of 19 of the 24 low SNR sources.

10. Individual images of the remaining 42 sources revealed that a further two islands were also false detections, reducing the final number of sources within the COBRaS L-band All Source Catalogue to 40.

The purpose of much of the methodology outlined above was to eliminate the false detections from SEAC. These were numerous ($\sim$ 65%) due to the low value of the flood fill threshold, $\sigma_f$, used at just three $\times$ the local RMS of the image. Moreover, the majority of

the false detections were found towards the outer regions of each wide field image where the noise level increases due to the response of the primary beam (see Section 2.4.3 for further details). For an image containing 19784 × 19784 pixels (i.e. $3.914 \times 10^8$) whose values can be represented by a Gaussian profile, $\sim$ 224 of these pixels will have a value larger than $5\sigma$. Furthermore $\sim 1.056 \times 10^6$ pixels will have a value large than 3 × the image RMS. Whilst the floodfill algorithm within SEAC is expected to limit the number of false detections it is clear that a certain number of false detections are expected when processing images containing such large numbers of pixels.

To highlight the sensitivity of the number of detected sources with the value of $\sigma_f$ used, SEAC was run again on each of the wide field images with values of $\sigma_s = 5.0$ and $\sigma_f = 4.0$. This resulted in a total number of 59 detections (in comparison to the 119 detections found using a value of $\sigma_f = 3.0$). The total number of individual sources detected (having removed the duplicates where a source was detected across multiple pointings) was 39 and three of these were still associated with artefacts within the immediate vicinity around a bright source. Thus the final number of sources was 36, all of which were also found in the $\sigma_s = 3.0$ SEAC runs. Therefore, despite the large number of false detections within the original $\sigma_s$=3.0 SEAC runs, a further four sources (at very low flux densities) were distinguished, in comparison to the $\sigma_s$=4.0 SEAC runs.

Since the wide field images have a larger noise level in comparison to those produced when imaging a small region of the sky (i.e. like those used to validate or eliminate the true or false sources detected from the wide field images with SEAC), it is possible that sources close to the noise level have been missed. Unfortunately, this cannot be investigated unless every region of the sky covered by the observations are studied in small sections at a time, or if wide field images of a better quality can be produced. It is expected that very few sources have been missed due to the larger noise level associated with the wide field images and only one case has been found thus far. The previously unclassified source of SBHW 109, originally identified by the Westerbork radio synthesis telescope survey of Cyg OB2 by Setia Gunawan et al. (2003), was not detected in the wide field image of pointing C. However, when a 512×512 pixel image, centred upon the coordinates of SBHW109 was produced, the source was detected to have two components at flux densities of 126±32 $\mu$Jy

and 274±49 $\mu$Jy respectively. This source has since been added to the all source catalogue, bringing the number of individual detections up to 41. Of these 41 objects found within the COBRaS L-band observations taken on 2014 Apr. 26th:

(i) 15 were detected within two or three of the seven COBRaS L-band pointings.

(ii) 5 were found to have a more than one component as detected with SEAC.

(iii) 25 were also found within the Apr. 11th observations.

The positions, integrated flux densities and their associated uncertainties, S/N and local RMS of each object detected within the COBRaS L-band observations can be found in Table 3.3. The errors on the positional information can be found in Table C.1 of Appendix C. They have been calculated from a combination of the error in measuring the weighted mean pixel position of each source (see Equation 3.2) and positional error inherent to the radio interferometer (see Appendix C). Where the object was also detected in the other observation epoch (Apr. 11th), its integrated flux density (and uncertainty), the local RMS and signal-to-noise ratio associated with that epoch is also given. Information regarding the extent of the source on the sky in relation to the size of the beam can also be found in Table C.1 of Appendix C. Table 3.3 is accompanied by the following information.

- *Column(1)*: the identification number given to the sources from this piece of work, sequentially increasing in order of right ascension (RA).
- *Column(2)*: the pointing within which the source information has been measured from. This is in all cases the pointing centre closest to the position of the source.
- *Columns(3)-(4)*: the J2000 positional information of each source as measured from the Apr. 26th observations. Each position has been calculated from the individual images of each source with SEAC by taking a weighted mean of the detected pixel area (see Section 3.1.1.3). The positions are all accurate to $< 0.032$ arcseconds in both RA and DEC, the positional errors have been calculated as described in Appendix C, and can be found by each individual source in Table C.1.
- *Columns(5)-(7)*: the integrated flux density, $F_{26^{th}}$ and its associated error in $\mu$Jy, the $1\sigma$ RMS noise level, $\sigma_{26^{th}}$, around the immediate vicinity of the source (within

an 4" × 4" box surrounding the centre of the source) and the signal-to-noise ratio ($F_{26^{th}}/\sigma_{26^{th}}$) of each of the sources detected at 21cm from the Apr. 26th observations.

- *Columns(8)-(10)*: the same as for columns (5) – (7) but in relation to the objects as detected within the 21cm observations taken on Apr. 11th.

Following Table 3.3, a contour and greyscale plot of each detected source is presented. Each plot was created using the AIPS task KNTR. The 25 sources which appear within both observation epochs are first presented in Figure 3.10, where the left and right columns represent those observed within the Apr. 26th and Apr. 11th observations respectively. The contour and greyscale plots of the 16 sources detected solely in the observations taken on 2014 Apr. 26th are then given in Figure 3.12.

TABLE 3.3: The COBRaS L-band All Source Catalogue (CLASC).

| ID | Point | RA (J2000) | DEC (J2000) | $F_{26^{th}}$ ($\mu$Jy) | $\sigma_{26^{th}}$ ($\mu$Jy) | SNR $26^{th}$ | $F_{11^{th}}$ ($\mu$Jy) | $\sigma_{11^{th}}$ ($\mu$Jy) | SNR $11^{th}$ |
|---|---|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| 1 | E | 20 32 06.13 | 41 14 07.54 | 1574±166 | 24 | 64.3 | 1062±127 | 40 | 26.7 |
| 2 | E | 20 32 08.29 | 41 13 11.48 | 213±38 | 22 | 9.5 | ‡101±32 | 37 | 2.7 |
| 3 | E | 20 32 14.23 | 41 16 38.25 | 236±42 | 24 | 9.9 | ‡132±40 | 39 | 3.3 |
| 4 | E | 20 32 15.46 | 41 13 31.33 | 145±29 | 21 | 6.8 | †115±34 | 34 | 3.3 |
| 5 | G | 20 32 20.59 | 41 00 20.45 | 213±57 | 53 | 4.0 | - | 109 | - |
| 6a (Cyg 5) | E | 20 32 22.42 | 41 18 18.88 | 455±74 | 36 | 12.6 | - | 45 | - |
| 6b (Cyg 5) | E | 20 32 22.48 | 41 18 19.44 | 921±119 | 32 | 29.1 | 226±54 | 45 | 5.0 |
| 7 | G | 20 32 26.87 | 41 04 32.97 | 3664±375 | 32 | 114 | 2903±311 | 59 | 48.9 |
| 8 (A11) | E | 20 32 31.53 | 41 14 08.14 | 219±35 | 21 | 10.3 | - | 33 | - |
| 9 | B | 20 32 34.21 | 41 24 14.17 | 520±68 | 27 | 19.3 | - | 49 | - |
| 10 (SBHW 81) | E | 20 32 36.61 | 41 14 48.05 | 2529±265 | 27 | 94.5 | 1550±176 | 39 | 39.8 |
| 11a (SBHW 83) | B | 20 32 38.17 | 41 23 37.15 | 3591±422 | 47 | 76.8 | 900±147 | 56 | 16.1 |
| 11b (SBHW 83) | B | 20 32 38.18 | 41 23 38.67 | 908±183 | 60 | 15.1 | - | - | - |
| 11c (SBHW 83) | B | 20 32 38.20 | 41 23 39.36 | 1257±213 | 52 | 23.9 | - | - | - |
| 12 (Cyg 12) | E | 20 32 40.95 | 41 14 29.20 | 1050±119 | 25 | 42.8 | 610±85 | 39 | 15.6 |
| 13 | B | 20 32 42.09 | 41 24 00.24 | †97±29 | 27 | 3.62 | - | 45 | - |
| 14 | G | 20 32 48.34 | 41 03 03.98 | †98±29 | 26 | 3.8 | - | 48 | - |
| 15 | D | 20 32 49.23 | 41 14 59.56 | †127±36 | 30 | 4.2 | - | 54 | - |
| 16 | D | 20 32 49.93 | 41 14 48.21 | ‡100±29 | 28 | 3.5 | - | 53 | - |
| 17 (SBHW 90) | G | 20 32 56.79 | 41 08 53.37 | 14661±1485 | 42 | 352.3 | 10798±1098 | 67 | 161.0 |
| 18 | G | 20 32 59.07 | 41 04 58.65 | 742±89 | 27 | 27.4 | 438±76 | 51 | 8.5 |
| 19 | D | 20 33 07.66 | 41 08 54.58 | 329±53 | 30 | 11.1 | - | 55 | - |
| 20 | G | 20 33 09.59 | 41 05 05.76 | 673±83 | 27 | 24.5 | ‡214±58 | 57 | 3.7 |
| 21 (Cyg 9) | D | 20 33 10.73 | 41 15 08.14 | 1307±139 | 25 | 52.4 | 791±129 | 47 | 17 |
| 22 (Cyg 8A) | A | 20 33 15.07 | 41 18 50.42 | 1083±122 | 34 | 32.1 | 2817±414 | 65 | 43.5 |
| 23 | F | 20 33 18.29 | 41 02 11.50 | 229±44 | 32 | 7.3 | ‡133±43 | 53 | 3.4 |
| 24a | A | 20 33 23.52 | 41 27 25.48 | 258±75 | 54 | 4.8 | - | 98 | - |
| 24b | A | 20 33 23.58 | 41 27 25.56 | 1444±188 | 46 | 31.5 | - | 98 | - |
| 25 | F | 20 33 23.59 | 41 09 17.45 | 245±44 | 28 | 8.7 | - | 50 | - |
| 26 | F | 20 33 26.99 | 41 08 53.27 | 2791±289 | 34 | 82.8 | 2281±247 | 53 | 43.1 |
| 27 | F | 20 33 32.09 | 41 05 57.65 | 299±45 | 23 | 12.9 | 208±44 | 37 | 5.6 |
| 28a | F | 20 33 33.79 | 40 59 42.63 | 327±69 | 43 | 7.5 | 302±81 | 64 | 4.7 |
| 28b | F | 20 33 33.86 | 40 59 43.27 | 1092±143 | 39 | 28.3 | - | 54 | - |
| 29 | A | 20 33 36.53 | 41 21 33.54 | ‡67±20 | 21 | 3.2 | - | 48 | - |
| 30 | F | 20 33 44.69 | 41 05 46.11 | 631±82 | 25 | 25.3 | 208±45 | 39 | 5.4 |
| 31 | F | 20 33 47.37 | 41 04 56.80 | 140±33 | 26 | 5.4 | - | 44 | - |
| 32 | A | 20 33 51.98 | 41 21 51.56 | 1890±214 | 27 | 71.2 | 483±98 | 58 | 8.3 |
| 33a (SBHW 109) | C | 20 33 52.27 | 41 15 46.83 | 126±32 | 27 | 4.9 | - | 48 | - |
| 33b (SBHW 109) | C | 20 33 52.42 | 41 15 42.08 | 274±49 | 25 | 11.3 | - | 48 | - |
| 34 (SBHW 110) | F | 20 33 55.48 | 41 02 53.45 | 3921±403 | 41 | 96 | 3521±375 | 64 | 55 |
| 35 (SBHW 112) | C | 20 33 58.33 | 41 09 14.55 | 2948±303 | 29 | 101.1 | 1424±187 | 63 | 22.6 |
| 36 | F | 20 34 01.47 | 41 03 41.96 | 1003±120 | 34 | 29.8 | 656±101 | 55 | 11.9 |
| 37 | A | 20 34 02.42 | 41 24 57.38 | 1472±174 | 41 | 35.9 | 881±155 | 92 | 9.5 |
| 38 | C | 20 34 06.76 | 41 16 00.56 | 405±58 | 25 | 16.4 | ‡177±50 | 49 | 3.6 |
| 39 | C | 20 34 07.30 | 41 07 24.60 | 151±38 | 34 | 4.4 | - | 67 | - |
| 40 | C | 20 34 18.14 | 41 15 44.72 | 1627±184 | 29 | 55.7 | - | 57 | - |
| 41 | C | 20 34 32.71 | 41 11 16.93 | 1297±158 | 38 | 34.4 | - | 75 | - |

Notes: where possible ID numbers are followed in brackets with an additional identifier (i.e. the common reference for that object used throughout this thesis). SBHW: Setia Gunawan et al. (2003); Cyg: Schulte (1958); A: Comerón et al. (2002).

† these sources are faint detections and their flux density has been measured with SEAC using threshold values, $\sigma_s = 4.0$ and $\sigma_f = 2.6$.

‡ these sources are very faint detections meaning not only does the same apply to them as those marked with a †, but the detected pixel area is smaller than the beam size, i.e. $\theta_{beam} > Ss > 0.75\,\theta_{beam}$, where $Ss$ is the source size in pixels and $\theta_{beam}$ is the size of the beam in pixels.

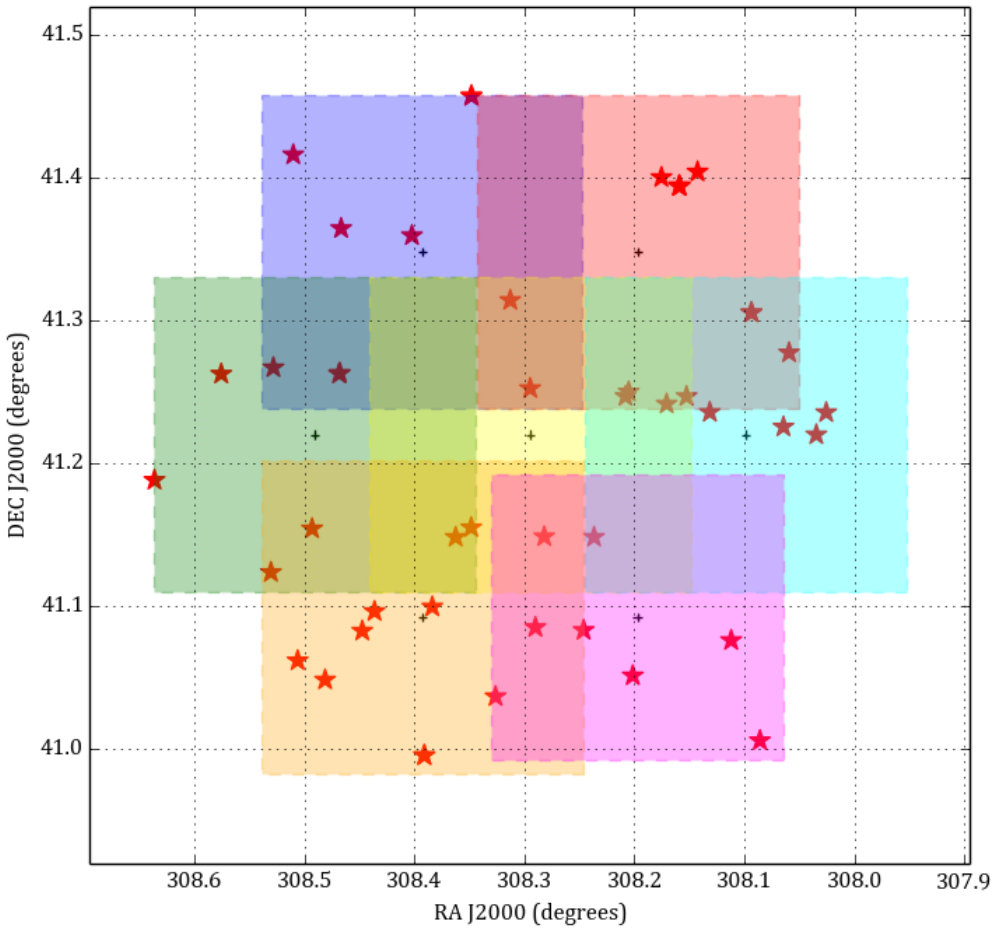
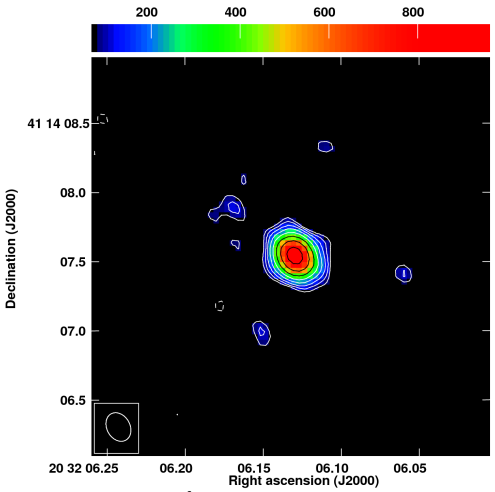
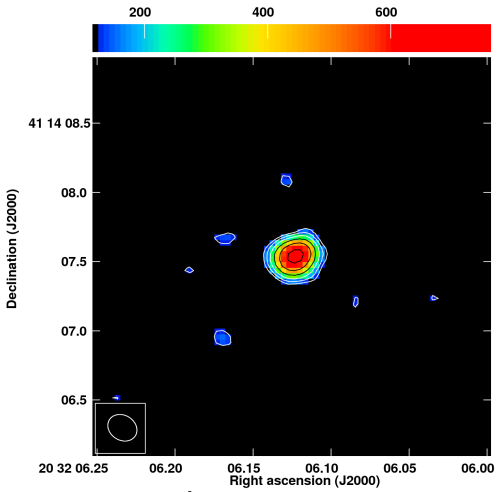


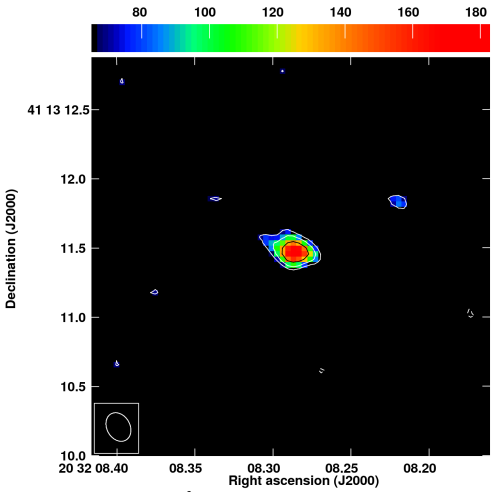

FIGURE 3.9: The positions of the 41 sources (shown in red) derived from the CLASC overlaid onto the areas as marked by the seven COBRaS L-band pointings. The central position of each pointing is marked by a black cross.



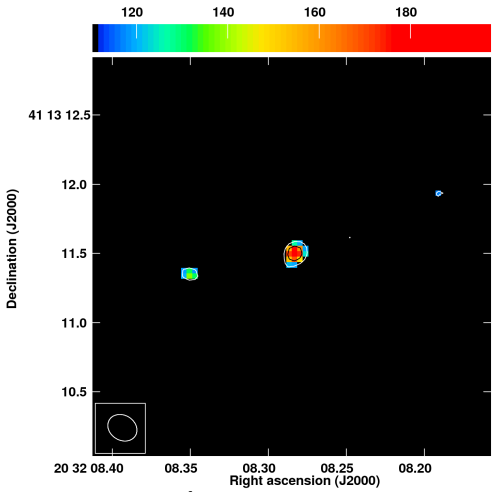

(A) **ID#1** ($26^{th}$): RMS = 24$\mu$Jy/beam, CSR: 72 - 956 $\mu$Jy/beam.



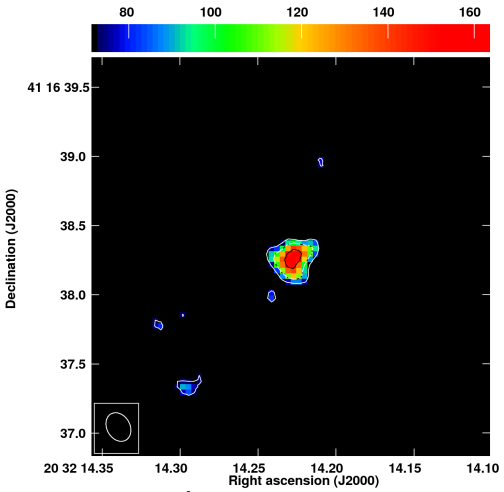

(B) **ID#1** ($11^{th}$): RMS = 40$\mu$Jy/beam, CSR: 120 - 758 $\mu$Jy/beam.



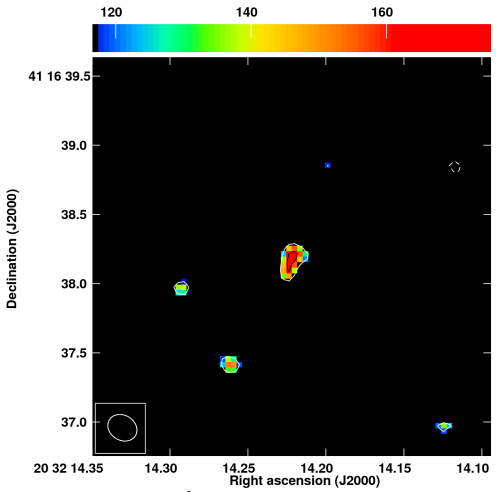

(C) **ID#2** ($26^{th}$): RMS = 22$\mu$Jy/beam, CSR: 66 - 182 $\mu$Jy/beam.




(D) **ID#2** ($11^{th}$): RMS = 37$\mu$Jy/beam, CSR: 111 - 197 $\mu$Jy/beam.

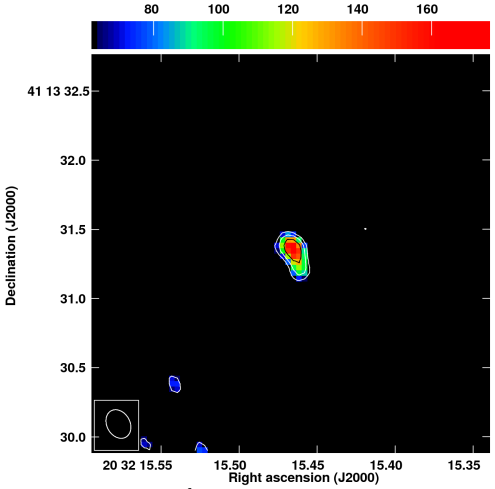



(E) **ID#3** ($26^{th}$): RMS = 24$\mu$Jy/beam, CSR: 72 - 163 $\mu$Jy/beam.

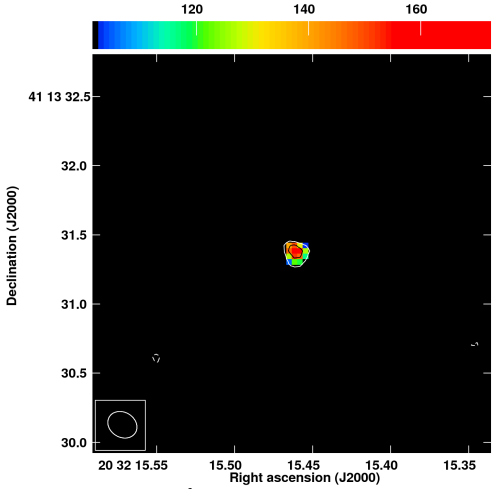



(F) **ID#3** ($11^{th}$): RMS = 39$\mu$Jy/beam, CSR: 117 - 175 $\mu$Jy/beam.

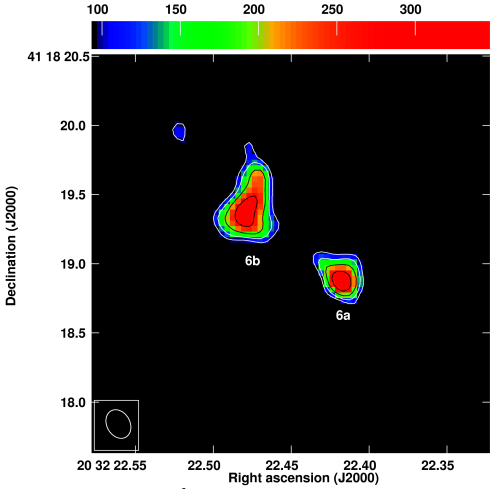
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



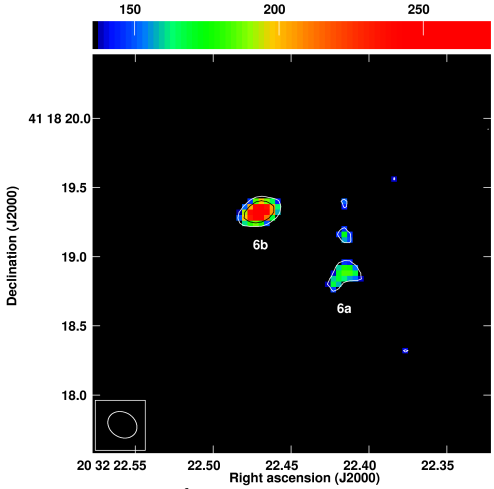

(G) **ID#4** ($26^{th}$): RMS = 21$\mu$Jy/beam, CSR: 63 - 176 $\mu$Jy/beam.



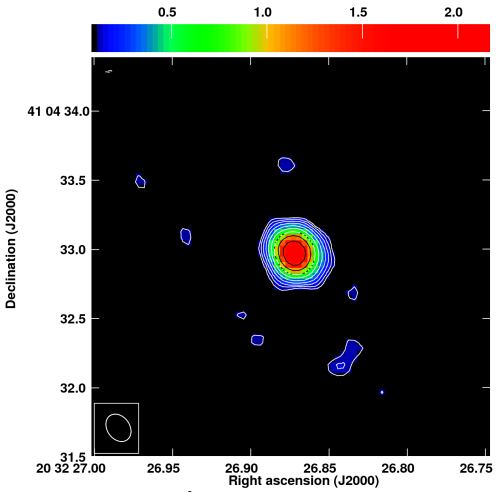

(H) **ID#4** ($11^{th}$): RMS = 34$\mu$Jy/beam, CSR: 102 - 172 $\mu$Jy/beam.



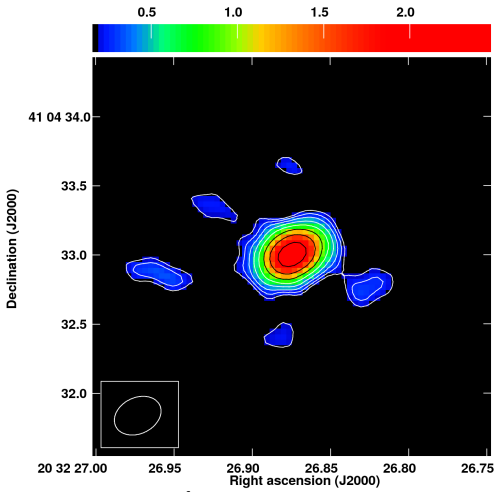

(I) **ID#6** ($26^{th}$): RMS = 32$\mu$Jy/beam, CSR: 95 - 346 $\mu$Jy/beam.




(J) **ID#6** ($11^{th}$): RMS = 46$\mu$Jy/beam, CSR: 137 - 272 $\mu$Jy/beam.

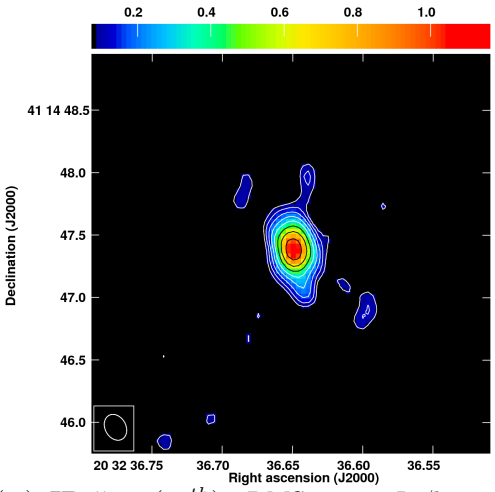



(K) **ID#7** ($26^{th}$): RMS = 32$\mu$Jy/beam, CSR: 96 - 2153 $\mu$Jy/beam.

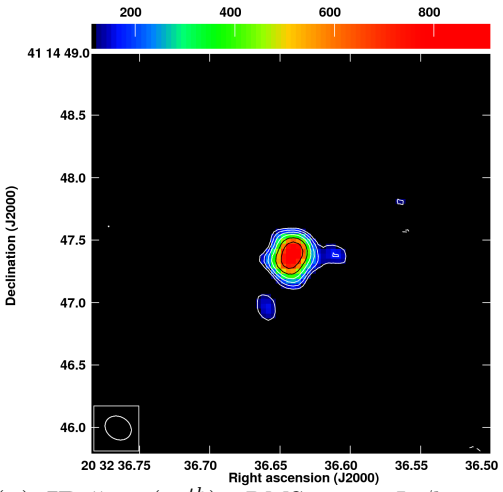



(L) **ID#7** ($11^{th}$): RMS = 59$\mu$Jy/beam, CSR: 177 - 2453 $\mu$Jy/beam.

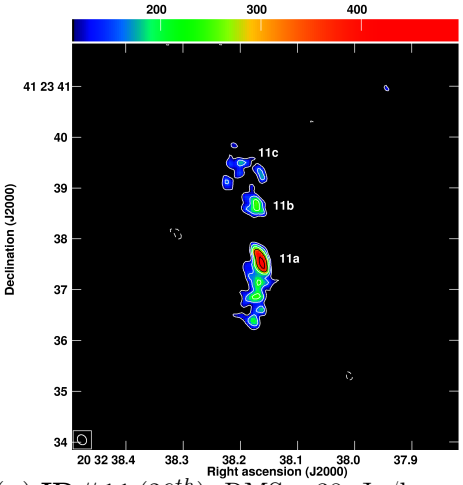
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



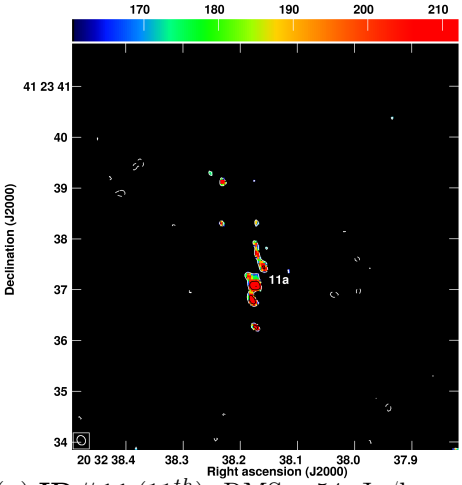

(M) **ID#10** ($26^{th}$): RMS = 27$\mu$Jy/beam, CSR: 80 - 1156 $\mu$Jy/beam.



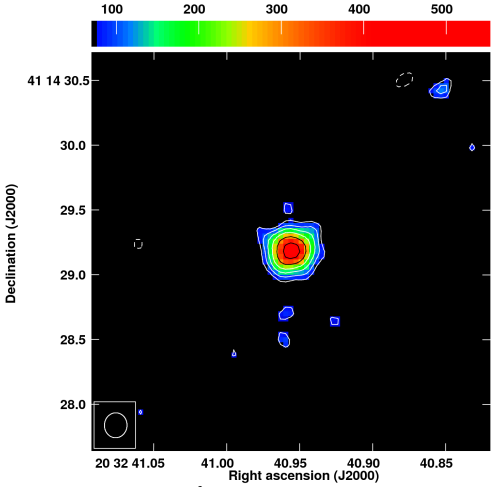

(N) **ID#10** ($11^{th}$): RMS = 39$\mu$Jy/beam, CSR: 117 - 908 $\mu$Jy/beam.



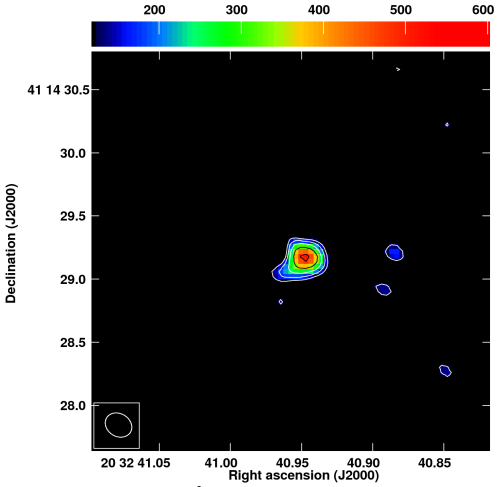

(O) **ID#11** ($26^{th}$): RMS = 38$\mu$Jy/beam, CSR: 113 - 469 $\mu$Jy/beam.




(P) **ID#11** ($11^{th}$): RMS = 54$\mu$Jy/beam, CSR: 161 - 212 $\mu$Jy/beam.

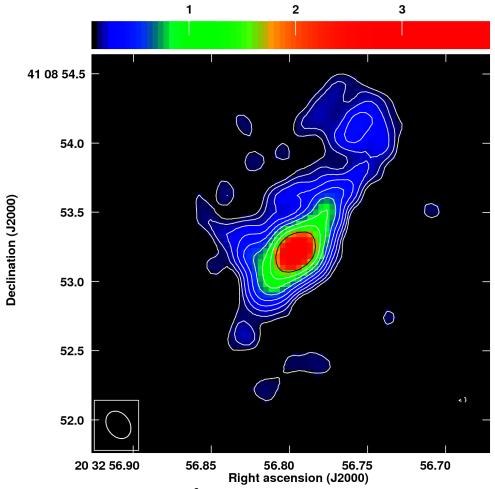



(Q) **ID#12** ($26^{th}$): RMS = 24$\mu$Jy/beam, CSR: 72 - 518 $\mu$Jy/beam.

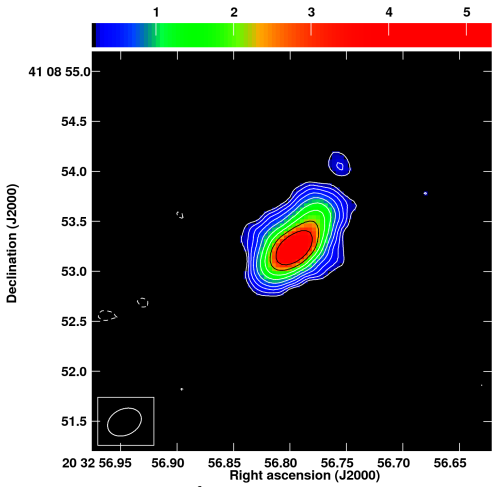



(R) **ID#12** ($11^{th}$): RMS = 40$\mu$Jy/beam, CSR: 120 - 510 $\mu$Jy/beam.

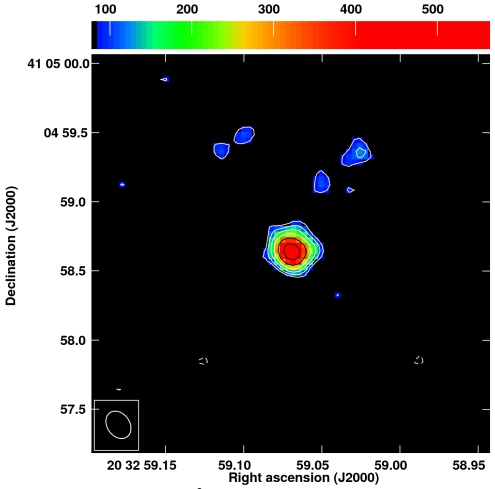
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



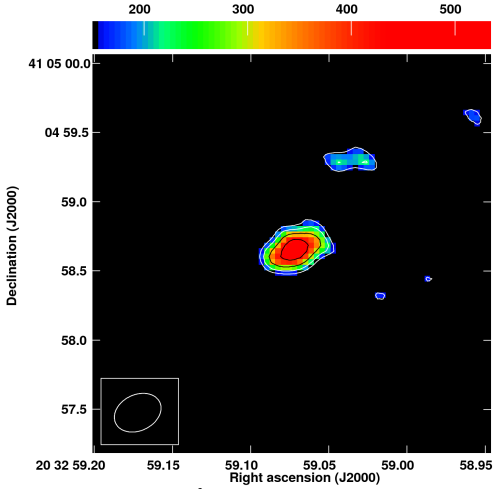

(S) **ID#17** ($26^{th}$): RMS = 42$\mu$Jy/beam, CSR: 111 - 3800 $\mu$Jy/beam.



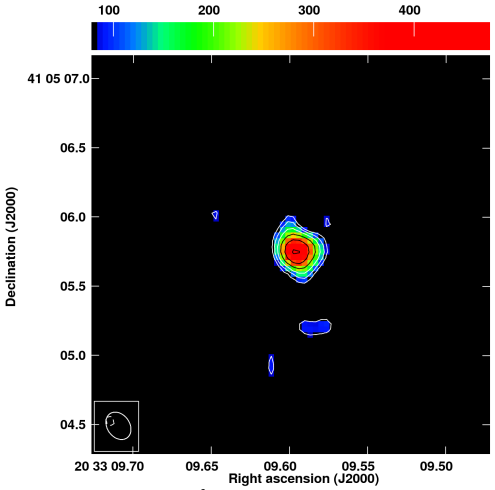

(T) **ID#17** ($11^{th}$): RMS = 67$\mu$Jy/beam, CSR: 200 - 5300 $\mu$Jy/beam.



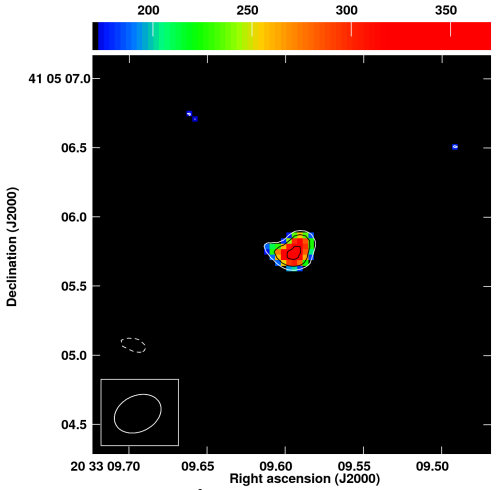

(U) **ID#18** ($26^{th}$): RMS = 27$\mu$Jy/beam, CSR: 81 - 561 $\mu$Jy/beam.




(V) **ID#18** ($11^{th}$): RMS = 51$\mu$Jy/beam, CSR: 153 - 532 $\mu$Jy/beam.

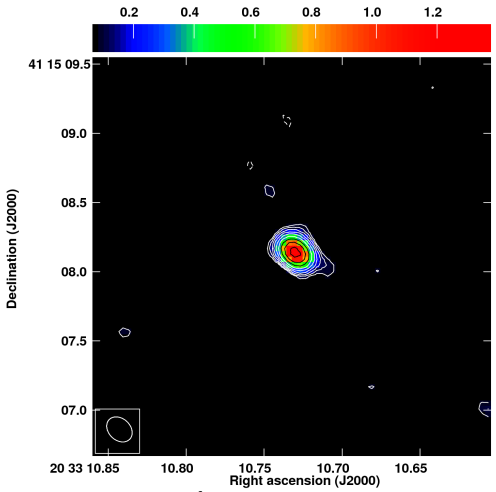



(W) **ID#20** ($26^{th}$): RMS = 27$\mu$Jy/beam, CSR: 81 - 473 $\mu$Jy/beam.

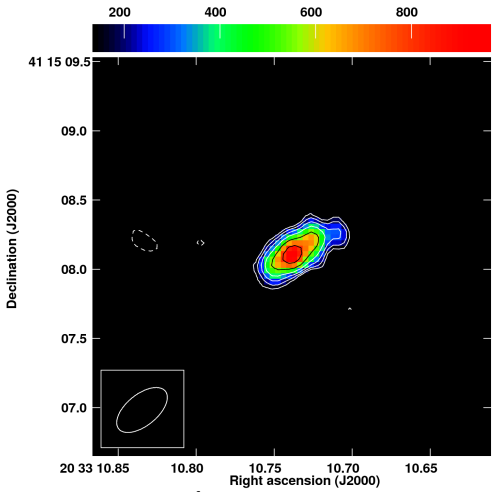



(X) **ID#20** ($11^{th}$): RMS = 57$\mu$Jy/beam, CSR: 171 - 369 $\mu$Jy/beam.

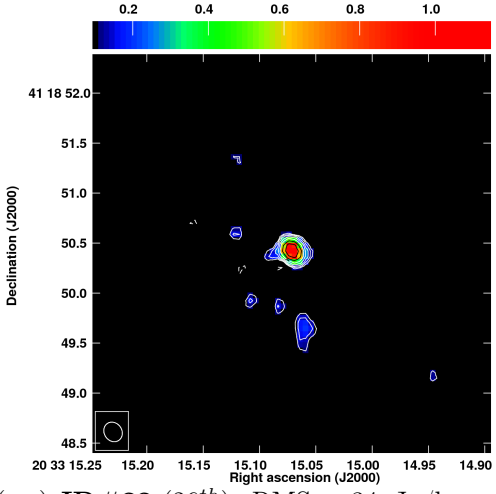
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



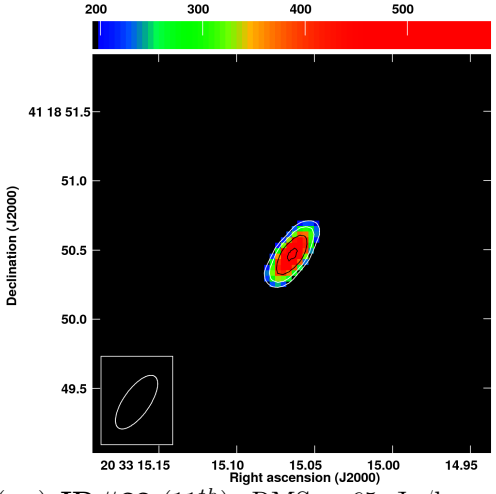

(Y) **ID#21** ($26^{th}$): RMS = 25$\mu$Jy/beam, CSR: 75 - 1368 $\mu$Jy/beam.



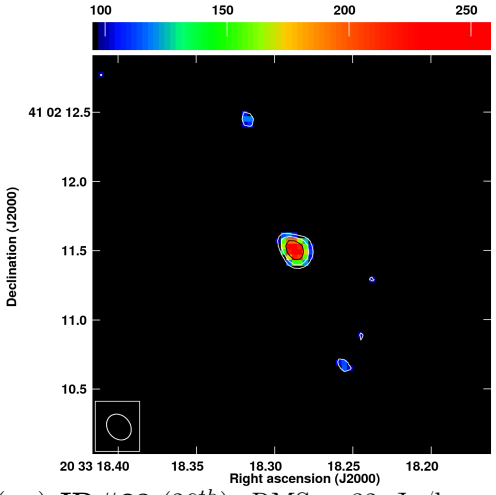

(Z) **ID#21** ($11^{th}$): RMS = 47$\mu$Jy/beam, CSR: 140 - 960 $\mu$Jy/beam.



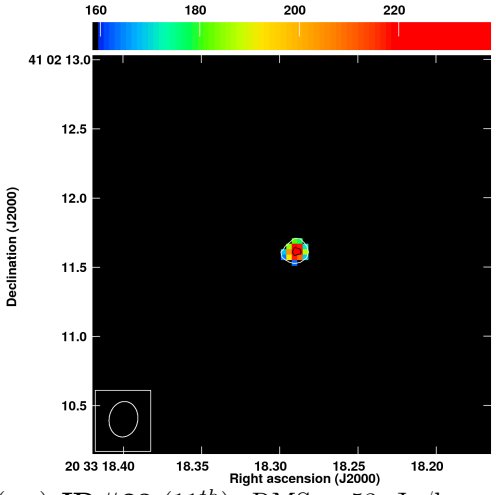

(AA) **ID#22** ($26^{th}$): RMS = 34$\mu$Jy/beam, CSR: 101 - 1140 $\mu$Jy/beam.




(AB) **ID#22** ($11^{th}$): RMS = 65$\mu$Jy/beam, CSR: 195 - 579 $\mu$Jy/beam.

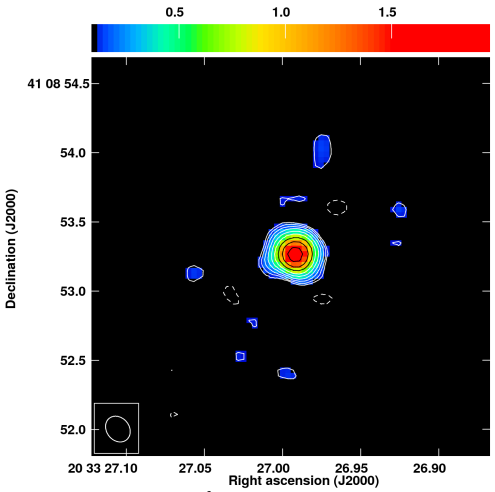



(AC) **ID#23** ($26^{th}$): RMS = 32$\mu$Jy/beam, CSR: 96 - 257 $\mu$Jy/beam.

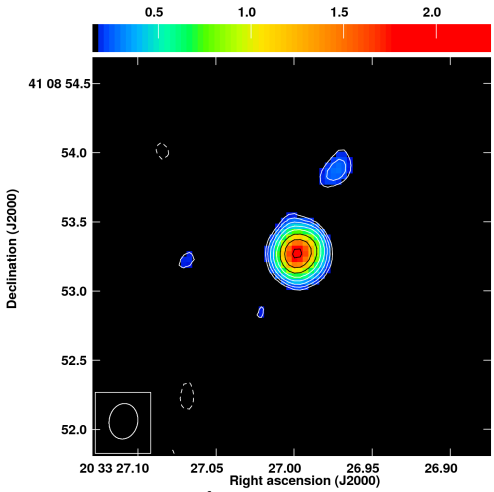



(AD) **ID#23** ($11^{th}$): RMS = 53$\mu$Jy/beam, CSR: 159 - 238 $\mu$Jy/beam.

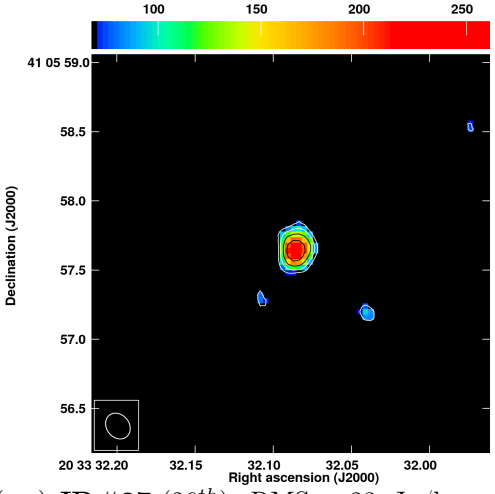
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



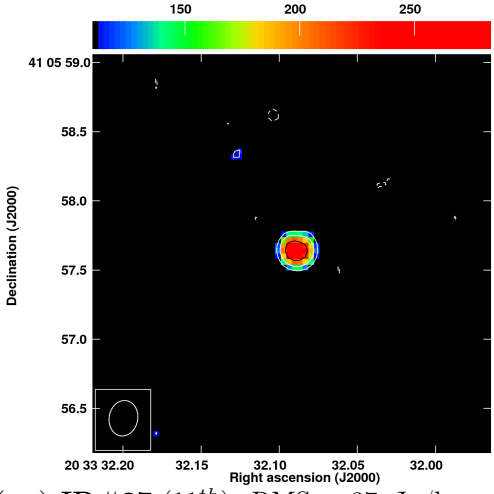

(AE) **ID#26** ($26^{th}$): RMS = 34$\mu$Jy/beam, CSR: 102 - 1950 $\mu$Jy/beam.



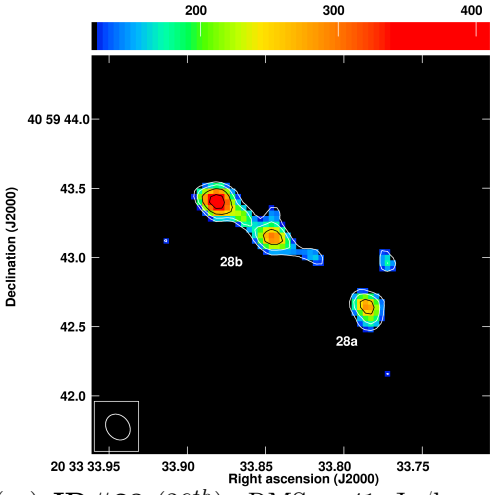

(AF) **ID#26** ($11^{th}$): RMS = 53$\mu$Jy/beam, CSR: 159 - 2281 $\mu$Jy/beam.



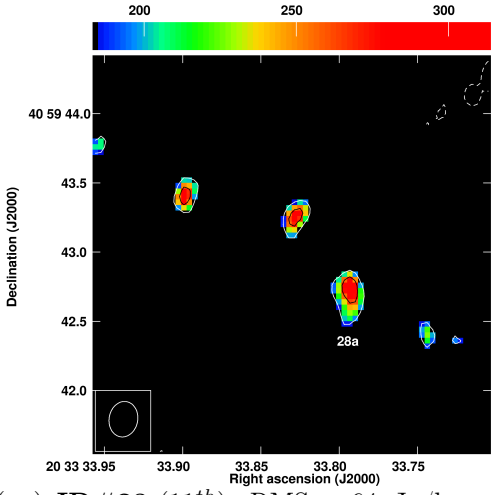

(AG) **ID#27** ($26^{th}$): RMS = 23$\mu$Jy/beam, CSR: 69 - 260 $\mu$Jy/beam.




(AH) **ID#27** ($11^{th}$): RMS = 37$\mu$Jy/beam, CSR: 111 - 282 $\mu$Jy/beam.

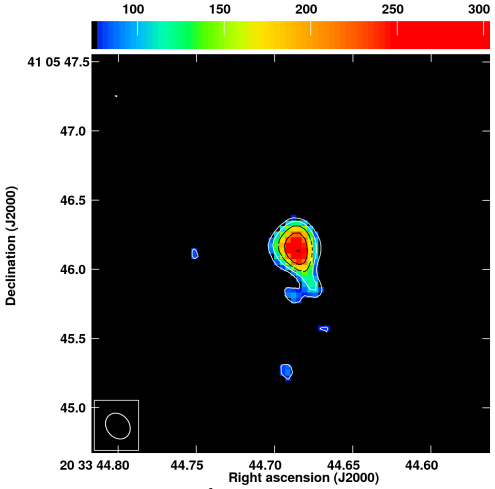



(AI) **ID#28** ($26^{th}$): RMS = 41$\mu$Jy/beam, CSR: 123 - 408 $\mu$Jy/beam.

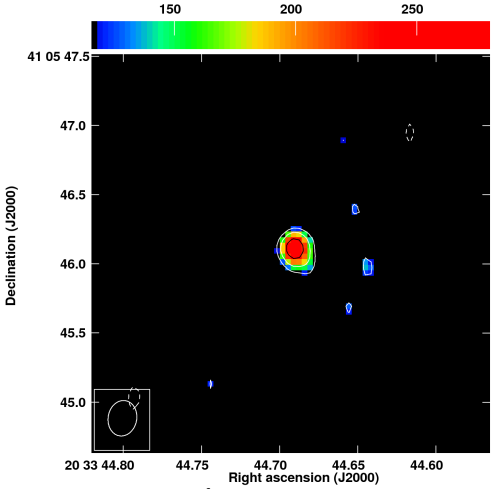



(AJ) **ID#28** ($11^{th}$): RMS = 64$\mu$Jy/beam, CSR: 184 - 313 $\mu$Jy/beam.

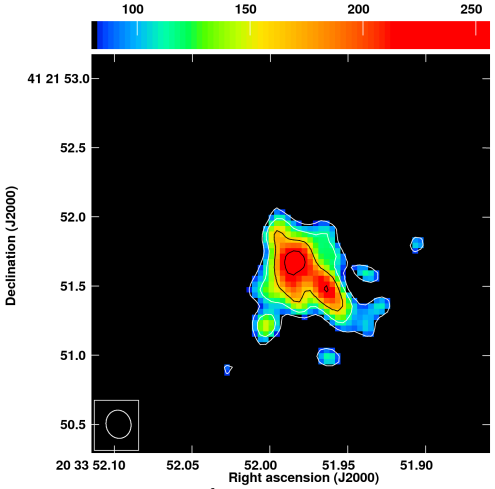
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



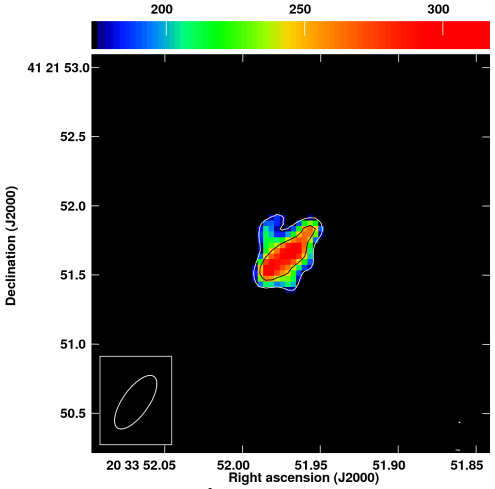

(AK) **ID#30** ($26^{th}$): RMS = 25$\mu$Jy/beam, CSR: 75 - 302 $\mu$Jy/beam.



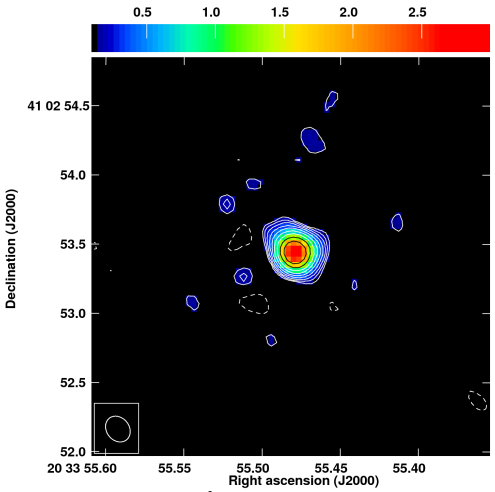

(AL) **ID#30** ($11^{th}$): RMS = 39$\mu$Jy/beam, CSR: 117 - 279 $\mu$Jy/beam.



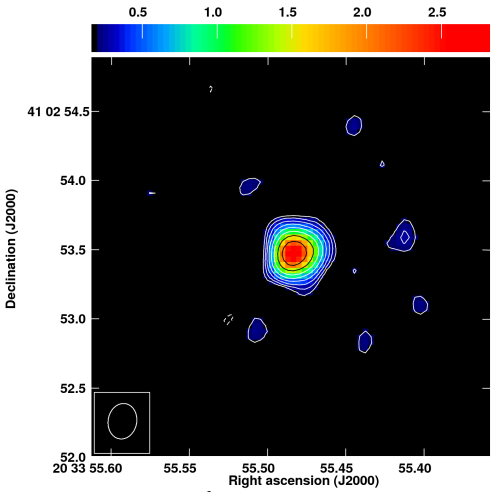

(AM) **ID#32** ($26^{th}$): RMS = 27$\mu$Jy/beam, CSR: 81 - 255 $\mu$Jy/beam.




(AN) **ID#32** ($11^{th}$): RMS = 58$\mu$Jy/beam, CSR: 174 - 316 $\mu$Jy/beam.

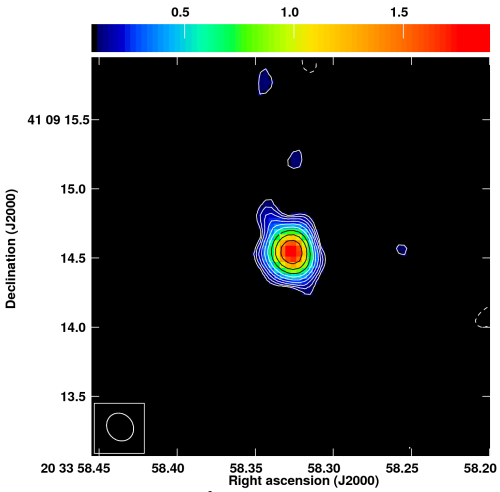



(AO) **ID#34** ($26^{th}$): RMS = 40$\mu$Jy/beam, CSR: 120 - 2975 $\mu$Jy/beam.

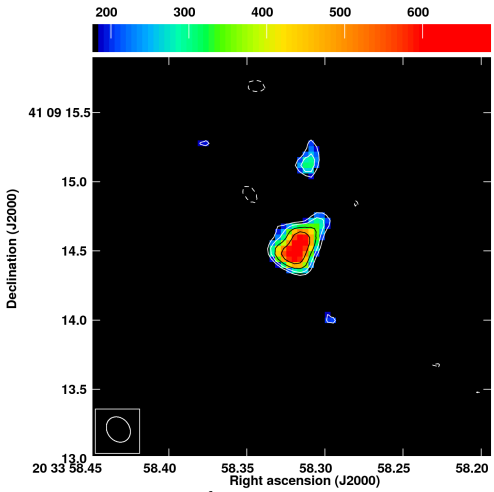



(AP) **ID#34** ($11^{th}$): RMS = 60$\mu$Jy/beam, CSR: 180 - 2805 $\mu$Jy/beam.

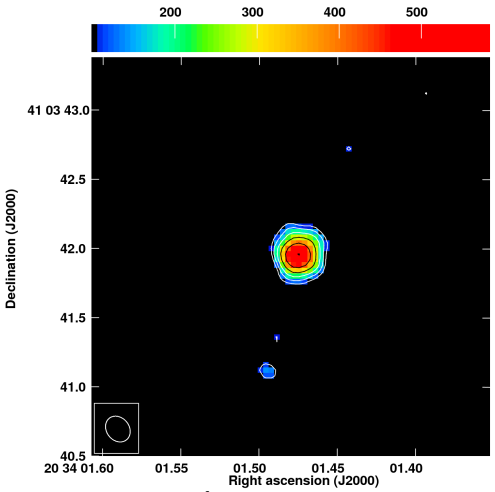
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



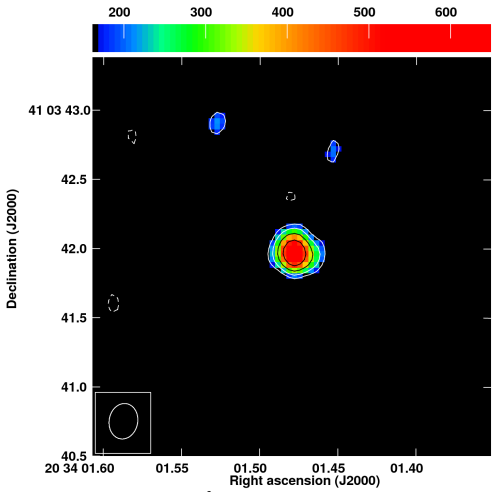

(AQ) **ID#35** ($26^{th}$): RMS = 29$\mu$Jy/beam, CSR: 87 - 1883 $\mu$Jy/beam.



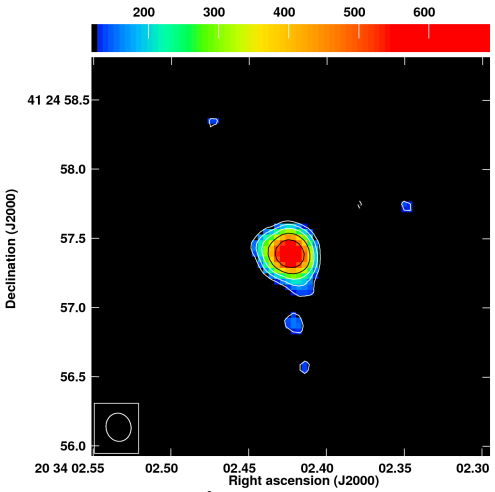

(AR) **ID#35** ($11^{th}$): RMS = 60$\mu$Jy/beam, CSR: 180 - 684 $\mu$Jy/beam.



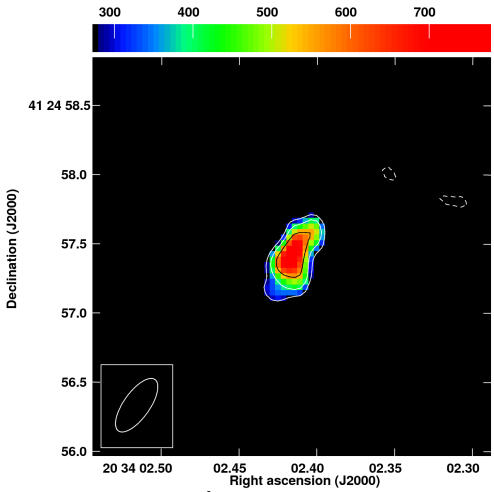

(AS) **ID#36** ($26^{th}$): RMS = 34$\mu$Jy/beam, CSR: 102 - 580 $\mu$Jy/beam.




(AT) **ID#36** ($11^{th}$): RMS = 55$\mu$Jy/beam, CSR: 165 - 646 $\mu$Jy/beam.

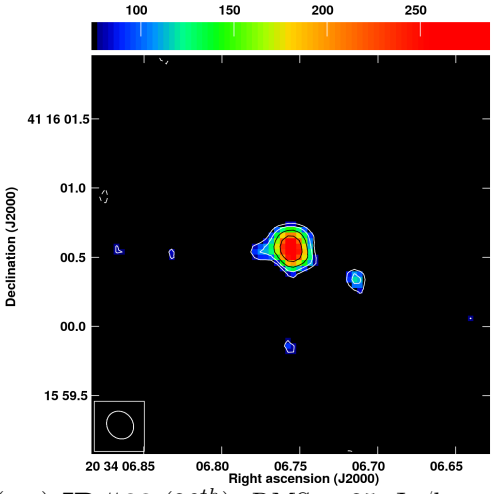



(AU) **ID#37** ($26^{th}$): RMS = 41$\mu$Jy/beam, CSR: 123 - 683 $\mu$Jy/beam.

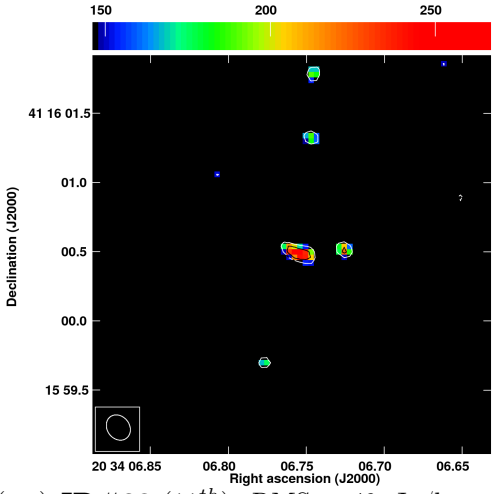



(AV) **ID#37** ($11^{th}$): RMS = 92$\mu$Jy/beam, CSR: 276 - 775 $\mu$Jy/beam.

FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.

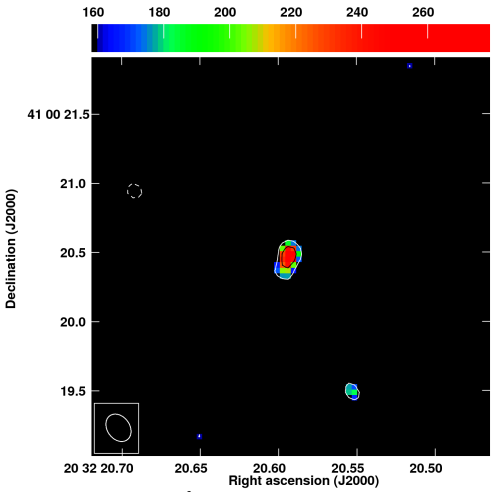



(AW) **ID#38** ($26^{th}$): RMS = 25$\mu$Jy/beam, CSR: 75 - 286 $\mu$Jy/beam.

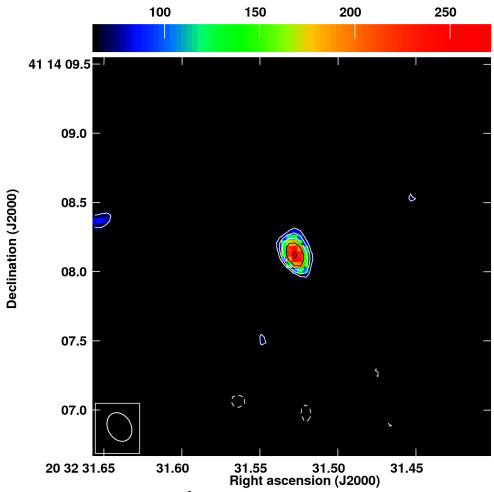



(AX) **ID#38** ($11^{th}$): RMS = 49$\mu$Jy/beam, CSR: 147 - 266 $\mu$Jy/beam.

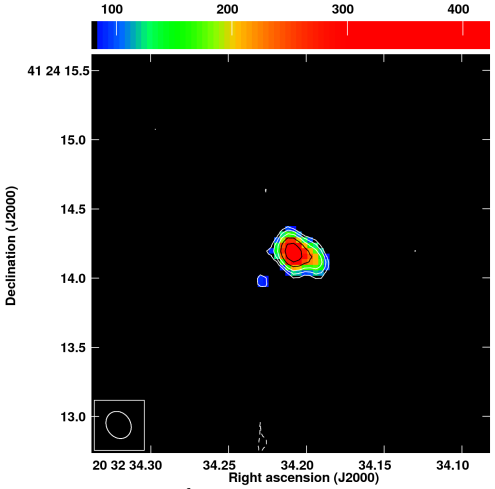
FIGURE 3.10: CLASC sources detected in both the Apr. 26th (left) and Apr. 11th (right) observation epochs. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



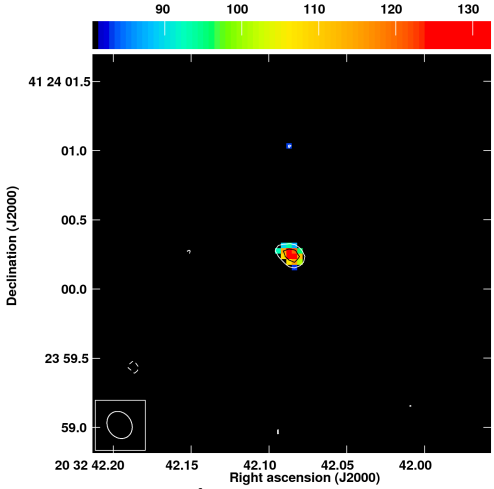

(A) **ID#5** ($26^{th}$): RMS = 53$\mu$Jy/beam, CSR: 159 - 278 $\mu$Jy/beam.



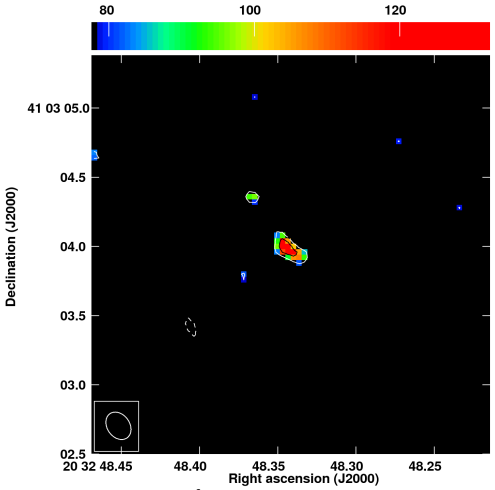

(B) **ID#8** ($26^{th}$): RMS = 21$\mu$Jy/beam, CSR: 66 - 270 $\mu$Jy/beam.



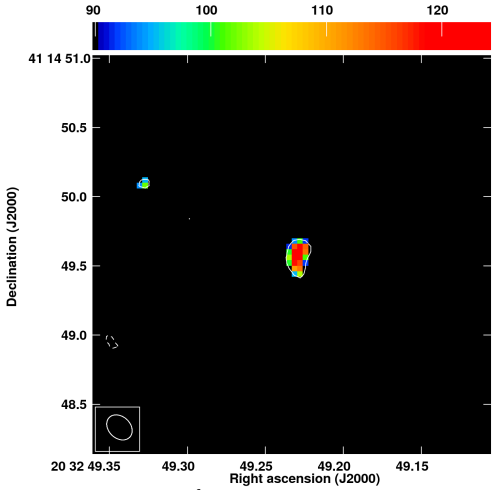

(C) **ID#9** ($26^{th}$): RMS = 27$\mu$Jy/beam, CSR: 81 - 421 $\mu$Jy/beam.




(D) **ID#13** ($26^{th}$): RMS = 27$\mu$Jy/beam, CSR: 81 - 132 $\mu$Jy/beam.

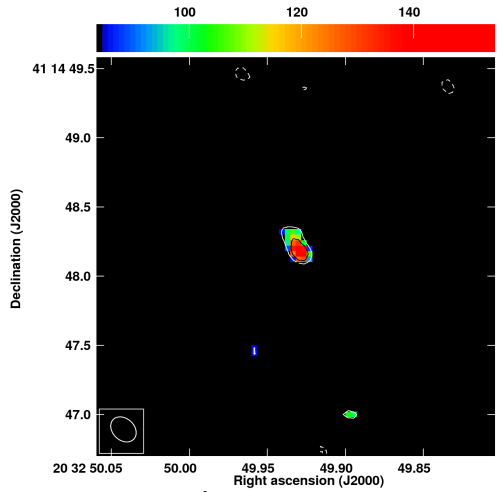



(E) **ID#14** ($26^{th}$): RMS = 26$\mu$Jy/beam, CSR: 78 - 132 $\mu$Jy/beam.

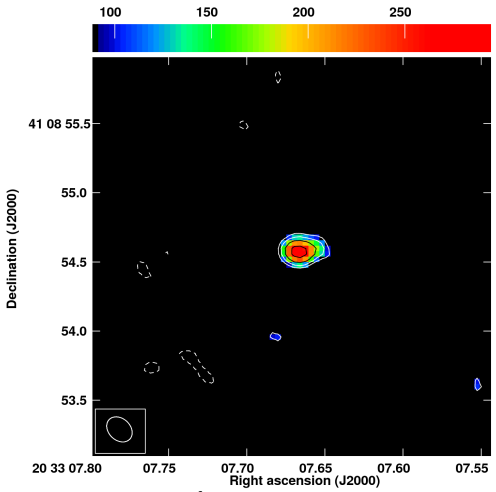



(F) **ID#15** ($26^{th}$): RMS = 30$\mu$Jy/beam, CSR: 90 - 124 $\mu$Jy/beam.

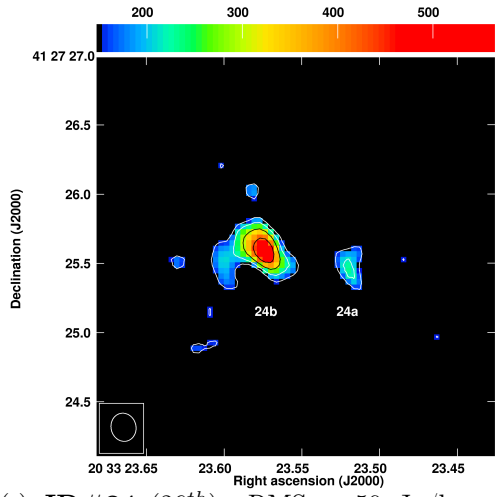
FIGURE 3.11: CLASC sources only detected in the Apr. 26th observations. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



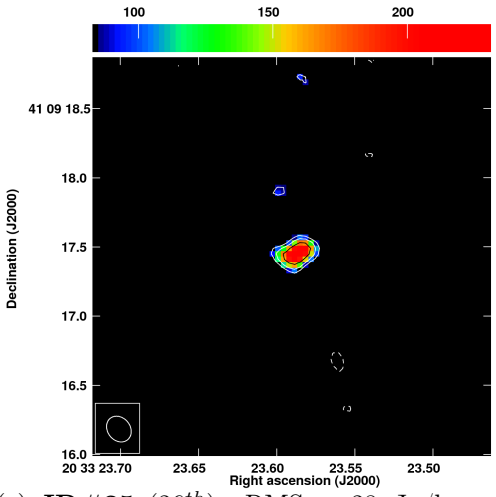

(G) **ID#16** ($26^{th}$): RMS = 28$\mu$Jy/beam, CSR: 84 - 154 $\mu$Jy/beam.



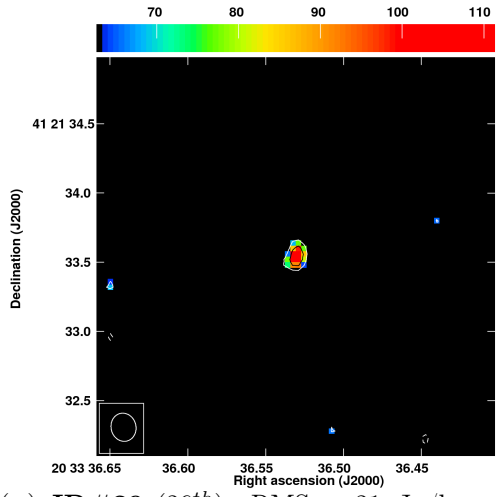

(H) **ID#19** ($26^{th}$): RMS = 30$\mu$Jy/beam, CSR: 90 - 293 $\mu$Jy/beam.



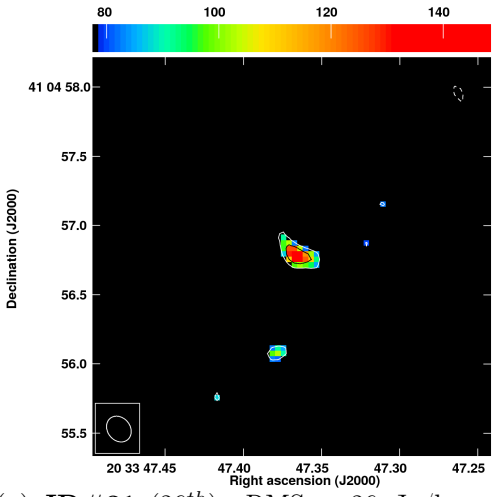

(I) **ID#24** ($26^{th}$): RMS = 50$\mu$Jy/beam, CSR: 50 - 562 $\mu$Jy/beam.




(J) **ID#25** ($26^{th}$): RMS = 28$\mu$Jy/beam, CSR: 84 - 230 $\mu$Jy/beam.

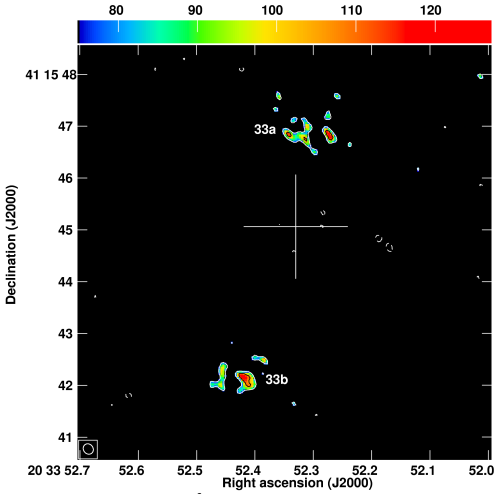



(K) **ID#29** ($26^{th}$): RMS = 21$\mu$Jy/beam, CSR: 63 - 111 $\mu$Jy/beam.

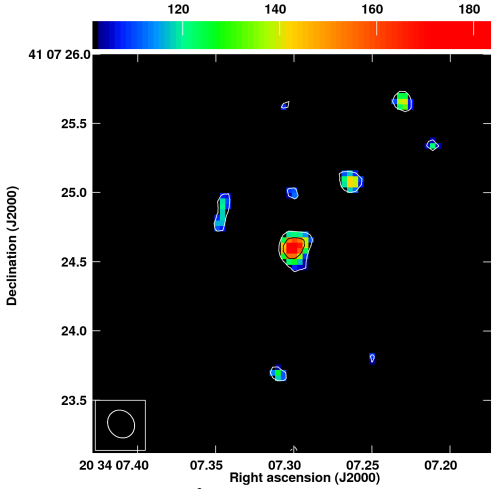



(L) **ID#31** ($26^{th}$): RMS = 26$\mu$Jy/beam, CSR: 78 - 148 $\mu$Jy/beam.

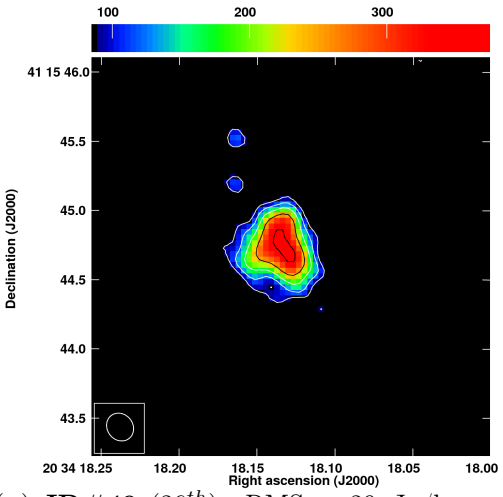
FIGURE 3.11: CLASC sources only detected in the Apr. 26th observations. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.



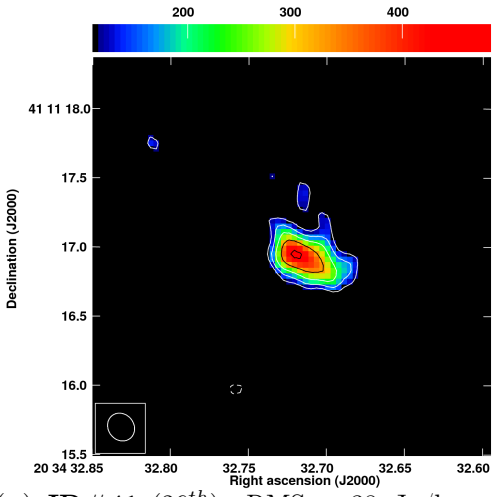

(A) **ID#33** ($26^{th}$): RMS = 25$\mu$Jy/beam, CSR: 75 - 127 $\mu$Jy/beam.




(B) **ID#39** ($26^{th}$): RMS = 34$\mu$Jy/beam, CSR: 102 - 183 $\mu$Jy/beam.




(C) **ID#40** ($26^{th}$): RMS = 29$\mu$Jy/beam, CSR: 87 - 373 $\mu$Jy/beam.




(D) **ID#41** ($26^{th}$): RMS = 38$\mu$Jy/beam, CSR: 114 - 480 $\mu$Jy/beam.

FIGURE 3.12: CLASC sources only detected in the Apr. 26th observations. Notes: CSR = colour scale range; RMS = root mean square of the local noise level.

### 3.3.2 Source identification

At this stage, each of the 41 detected objects were cross-correlated in turn with the Cyg OB2 super catalogue, the 2MASS catalogue and the G2012 catalogue. Each source was cross-matched to within an accuracy of 3.0 arcseconds in order to account for the large variety of observations that contributed to the Cyg OB2 catalogue. From the initial match, a total of 27 individual objects from CLASC were found to have positions within 3.0 arcseconds of at least one previously found object in the three cross-correlation catalogues. In order to validate these matches a method originally developed by De Ruiter, Willis, and Arp (1977), and also used by Setia Gunawan et al. (2003), was employed to find the likelihood ratio ($LR$) of a given match. The $LR$ is defined as the ratio between the probability distribution of position differences between a source and its identification, $dP_{id}$, and the probability distribution of the background objects, $dP_{bg}$ (De Ruiter, Willis, and Arp, 1977).

The probability distribution of the background objects is dependant on the properties of the observations from which the identification originates from, i.e. the number density of objects within the survey or observations $\rho(b)$, and the combined positional errors between the source and identification, $\sigma_\alpha$, $\sigma_\delta$.

$$dP_{bg} = 2\lambda r exp^{-\lambda r^2} dr \tag{3.5}$$

where,

$$\lambda = \pi \sigma_\alpha \sigma_\delta \rho(b) \tag{3.6}$$

$\lambda$ is the therefore the number of sources within an area corresponding to the size of the combined positional error in RA, $\sigma_\alpha = \sigma_{\alpha CLASC} + \sigma_{\alpha ID}$ and in DEC $\sigma_\delta = \sigma_{\delta CLASC} + \sigma_{\delta ID}$. The parameter $r$ is defined by De Ruiter, Willis, and Arp (1977), is the normalised

separation between a source and its identification:

$$r = \left( \frac{\Delta\alpha^2}{\sigma_\alpha^2} + \frac{\Delta\delta^2}{\sigma_\delta^2} \right)^{1/2} \tag{3.7}$$

where $\Delta\alpha$ and $\Delta\delta$ are the offsets in RA and DEC respectively, between the source and its identification. The probability therefore, of finding a true identification between $r$ and $r + dr$ can be written as:

$$dP_{id} = r\, exp^{-0.5r^2} dr \tag{3.8}$$

From which, the likelihood ratio of a given match can be written as the ratio between the probability of finding a true match and the probability distribution of background objects:

$$LR(r) = \frac{1}{2\lambda} exp\{[r^2(2\lambda - 1)/2]\} \tag{3.9}$$

Large values of $LR$ imply a more probable detection and a limiting value of $LR \approx 1$ is employed in order to claim a true identification.

The initial cross-correlation gave one or more identifications to 27 sources within CLASC. The well studied sources within the sample (Cyg OB2 #5, A11, Cyg OB2 #12, Cyg OB2 #9 and Cyg OB2 #8A) had numerous identifications from within the Cyg OB2 super catalogue and only the most recent identification from a given wavelength range (i.e. X-ray, IR, radio) were chosen, resulting in a total number of 65 identifications across the entire 41 sources from CLASC. In order to derive the source number densities, $\rho(b)$ and the positional uncertainties, the survey or observations associated with each identification were considered. Table 3.4 shows a summary of this information regarding the ten different observations or surveys contributing to the identifications found from the initial cross-correlation. The identifications of M2007 (Martí et al., 2007) and S2003 (Setia Gunawan et al., 2003) included positional uncertainties in RA and DEC for each individual source, as a result an average value across RA and DEC for each of their identifications is given in Table 3.4.

TABLE 3.4: Information regarding the specific catalogues from which sources from the CLASC have been matched to. Note for both M2007 and S2003 the positional accuracies represent average values since their positional uncertainties are individual to each source within their catalogue.

| Catalogue Reference | Waveband | Source Density (object arcsec$^{-2}$) | Positional Accuracy (arcsec) | Reference |
|---|---|---|---|---|
| M2007 | Radio | $2.29\times10^{-5}$ | ∼0.45 | 1 |
| S2009 | IR (V-band) | $2.21\times10^{-7}$ | 1.00 | 2 |
| R2011 | X-ray | $1.98\times10^{-4}$ | 0.63 | 3 |
| S2003 | Radio | $4.59\times10^{-6}$ | ∼1.07 | 4 |
| 2MASS | IR (JHK) | $4.16\times10^{-3}$ | 0.5 | 5 |
| M2005 | Radio | $2.22\times10^{-4}$ | 0.18 | 6 |
| W2014 | X-ray | $7.01\times10^{-4}$ | 0.6 | 7 |
| W2009 | X-ray | $8.15\times10^{-4}$ | 0.5 | 8 |
| G2012 | IR (RIZ) | $1.06\times10^{-2}$ | 0.13 | 9 |
| G2013⋆ | IR | $1.42\times10^{-4}$ | 1.2 | 10 |

**1** Martí et al. (2007); **2** Skiff (2009); **3** Rauw (2011); **4** Setia Gunawan et al. (2003); **5** Cutri et al. (2003); **6** Martí et al. (2005); **7** Wright et al. (2014); **8** Wright and Drake (2009); **9** Guarcello et al. (2012); **10** Guarcello et al. (2013).
⋆The only identification from G2013 (source ID: 38) was observed using the Spitzer Space Telescope.

Having calculated the likelihood ratio for each of the 65 possible identifications, those with a value of $LR \lesssim 1$ were discarded. Note that source ID: 9 was identified to a source from Martí et al. (2007) with a $LR = 0.8$ and was not discarded due to its close proximity to the limiting value of $LR$. This limiting value of $LR \approx 1$ was also used in the Westerbork radio survey by Setia Gunawan et al. (2003) and represents a good compromise between claiming a false identification and missing a true identification. Note also that all sources originally identified to those from G2012 (Guarcello et al., 2012) have been discarded since for these identifications, $LR << 1$, due to a combination of the large number density of the survey and small position errors (see Table 3.4).

After discarding all identifications with a $LR \lesssim 1$, a total of 34 remained, corresponding to a total number of true identifications for 19 of the 41 sources from CLASC. These are shown in Table 3.5.

- *Column(1)*: the ID number of each source from CLASC as given in Table 3.3.
- *Column(2)*: if any common identifier exists then this is listed, the references for which are as follows: 'Cyg #': Schulte (1958); 'A11': Comerón et al. (2002); 'SBHW

#': Setia Gunawan et al. (2003).

- *Column(3)*: the offset between the position of the CLASC source and that of the identification in arcseconds.
- *Columns(4)-(6)*: the calculated values of $\lambda$, $r$ and $LR$ as given by Equations 3.6, 3.7 and 3.9.
- *Column(7)*: the reference to the observations or survey from which each identification is taken (see Table 3.4).
- *Column(8)*: any additional information regarding the flux density, magnitude or X-ray flux of each identification.

### 3.3.3 Source counts

A source count was performed on each of the 41 objects detected in Table 3.3. This was carried out by first binning each of the 41 sources by their flux densities (according to that derived from the Apr. 26th observations). The flux density bins were defined from first starting at the lowest flux density source (ID #29 with $F_{21cm} = 67 \pm 20$ $\mu$Jy) and in incrementing bins of 0.5 in *log* space (i.e. $\sqrt[4]{2}$, as was also used by Hopkins et al. 1998), which provided a good resolution whilst maintaining a suitable number of sources within each bin. The final bin however, was made much larger to incorporate source ID #17, the highest flux density source in the sample at $F_{21cm} = 14.66 \pm 1.48$ mJy. Each of the bins are shown in Table 3.6 alongside the number of sources within them, $N_\nu$. This represents the number of sources within a given flux density range observed in a $\sim 0.48$ deg$^2$ area of the sky (i.e. that covered by the COBRaS L-band observations). This number is then converted into the number of sources per steradian, and divided by the central flux density of the bin to obtain the differential source count $dN/dS$, which is normalised to those expected in a Euclidean geometry (i.e. an isotropic, static universe with a non-evolving population) by dividing by $S^{-5/2}$. This final value is shown in the last column of Table 3.6.

TABLE 3.5: Source identification information.

| ID | Common ID | Offset (") | $\lambda$ | $r$ | $LR$ | Catalogue Ref | Flux Density/Mag/X-ray Flux |
|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 2 | - | 1.183 | 1.79E-005 | 2.5 | 1343.3 | M2007 | $F_{86cm}$ = 0.96 mJy |
| 6 | Cyg 5 | 0.108 | 6.94E-007 | 0.1 | $>10^3$ | S2009 | $V = 9.3$ |
| | | 2.391 | 2.47E-004 | 3.8 | 1.5 | R2011 | |
| | | 2.207 | 3.45E-005 | 1.3 | $>10^3$ | S2003 | $F_{21cm}$ = 3.6 mJy |
| | | 0.210 | 3.27E-003 | 0.4 | 140.2 | 2MASS | $J = 5.2, H = 4.7, K = 4.3$ |
| 7 | - | 0.144 | 2.13E-005 | 0.9 | $>10^3$ | M2005 | $F_{6cm}$=1.9 mJy, $F_{21cm}$=3.3 mJy |
| 8 | A11 | 0.200 | 7.93E-004 | 0.3 | 596.2 | W2014 | $XR_f = 5.08 \times 10^{-13}$ erg cm$^{-2}$ s$^{-1}$ |
| | | 0.250 | 3.26E-003 | 0.5 | 135.0 | 2MASS | $J = 7.8, H = 7.1, K = 6.7$ |
| 9 | - | 2.009 | 1.69E-005 | 4.6 | 0.8 | M2007 | $F_{86cm}$ = 1.45 mJy |
| | | 2.009 | 7.93E-004 | 3.3 | 2.3 | W2014 | $XR_f = 4.44 \times 10^{-15}$ erg cm$^{-2}$ s$^{-1}$ |
| 10 | SBHW 81 | 0.867 | 3.18E-005 | 0.6 | $>10^3$ | S2003 | $F_{21cm}$ = 3.6 mJy |
| 11 | SBHW 83 | 2.096 | 6.43E-006 | 2.9 | $>10^3$ | S2003 | $F_{21cm}$=8.4 mJy, $F_{49cm}$=62 mJy |
| 12 | Cyg 12 | 1.568 | 9.13E-006 | 4.6 | 1.6 | M2007 | $F_{86cm}$ = 0.93 mJy |
| | | 0.115 | 6.94E-007 | 0.1 | $>10^3$ | S2009 | $V = 11.5$ |
| | | 0.101 | 3.26E-003 | 0.2 | 149.9 | 2MASS | $J = 4.7, H = 3.5, K = 2.7$ |
| 17 | SBHW 90 | 1.004 | 5.65E-006 | 1.5 | $>10^3$ | S2003 | $F_{21cm}$=14.7 mJy, $F_{49cm}$=54 mJy |
| 18 | - | 0.416 | 7.93E-004 | 0.7 | 496.1 | W2014 | $XR_f = 4.36 \times 10^{-15}$ erg cm$^{-2}$ s$^{-1}$ |
| 21 | Cyg 9 | 0.120 | 7.93E-004 | 0.2 | 618.0 | W2014 | $XR_f = 1.39 \times 10^{-12}$ erg cm$^{-2}$ s$^{-1}$ |
| | | 0.171 | 6.94E-007 | 0.2 | $>10^3$ | S2009 | $V = 11.0$ |
| | | 0.295 | 3.27E-003 | 0.6 | 128.7 | 2MASS | $J = 6.5, H = 5.9, K = 5.6$ |
| 22 | Cyg 8A | 0.134 | 6.40E-004 | 0.3 | 753.2 | W2009 | |
| | | 0.147 | 6.94E-007 | 0.1 | $>10^3$ | S2009 | $V = 9.1$ |
| | | 1.767 | 2.47E-004 | 2.8 | 39.9 | R2011 | |
| | | 0.204 | 3.27E-003 | 0.4 | 140.8 | 2MASS | $J = 6.1, H = 5.7, K = 5.5$ |
| 27 | - | 1.243 | 7.93E-004 | 2.1 | 74.0 | W2014 | $XR_f = 1.26 \times 10^{-14}$ erg cm$^{-2}$ s$^{-1}$ |
| | | 1.243 | 6.40E-004 | 2.5 | 35.7 | W2009 | |
| 33 | SBHW 109 | 1.956 | 3.45E-005 | 1.2 | $>10^3$ | S2003 | $F_{21cm}$ = 3.0 mJy |
| 34 | SBHW 110 | 2.916 | 1.44E-005 | 2.6 | $>10^3$ | S2003 | $F_{21cm}$ = 2.2 mJy |
| | | 0.146 | 7.93E-004 | 0.2 | 612.2 | W2014 | $XR_f = 1.01 \times 10^{-14}$ erg cm$^{-2}$ s$^{-1}$ |
| 35 | SBHW 112 | 1.326 | 3.31E-005 | 0.8 | $>10^3$ | S2003 | $F_{21cm}$ = 2.2 mJy |
| 38 | - | 0.285 | 6.43E-004 | 0.2 | 755.6 | G2013 | $3.6_{\mu m}$=14.4, $4.5_{\mu m}$=13.3, $5.8_{\mu m}$=12.4, $8.0_{\mu m}$=11.6, $24_{\mu m}$=7.1 |
| 39 | - | 0.761 | 7.94E-004 | 1.3 | 282.8 | W2014 | $XR_f = 4.78 \times 10^{-15}$ erg cm$^{-2}$ s$^{-1}$ |
| 40 | - | 1.469 | 3.27E-003 | 0.4 | 140.1 | 2MASS | $J = 16.2, H = 14.9, K = 14.0$ |
| 41 | - | 0.334 | 7.93E-004 | 0.6 | 539.9 | W2014 | $XR_f = 1.37 \times 10^{-14}$ erg cm$^{-2}$ s$^{-1}$ |

TABLE 3.6: Source counts from the CLASC. The uncertainty in the normalised differential source count dN/dS was derived assuming a Poisson process, i.e. $\sqrt{N}$.

| Flux Density Range ($\mu$Jy) | $N_\nu$ | $N_\nu$ (sr$^{-1}$) | dN/dS $\times S^{5/2}$ (sr$^{-1}$ Jy$^{1.5}$) |
|---|---|---|---|
| 67-110.46 | 4 | 56993 | 0.097±0.049 |
| 110.46-182.12 | 5 | 71241 | 0.257±0.115 |
| 182.12-300.27 | 9 | 128234 | 0.981±0.327 |
| 300.27-495.07 | 4 | 56993 | 0.923±0.461 |
| 495.07-816.23 | 4 | 56993 | 1.953±0.977 |
| 816.23-1345.73 | 9 | 128234 | 9.304±3.101 |
| 1345.73-2218.74 | 5 | 71241 | 10.943±4.894 |
| 2218.74-3658.08 | 4 | 56993 | 18.533±9.266 |
| 3658.08-14661 | 3 | 42745 | 31.193±18.009 |

For comparison, these differential source counts are plotted alongside those obtained from previous radio surveys at 21cm in Figure 3.13. These include those from the Westerbork continuum survey of Cyg OB2 by Setia Gunawan et al. (2003, blue triangles) and from a number of VLA observations presented in Condon and Mitchell (1984, yellow squares). For sources below 1 mJy, the e-MERLIN 21cm COBRaS observations have observed significantly lower counts than those from Condon and Mitchell (1984). This is likely due to the high resolution obtained with the e-MERLIN instrument meaning that these COBRaS observations are less sensitive to low surface brightness sources which are 'resolved' out. For sources of flux density larger than 1 mJy, the derived source counts from the CLASC are in much better agreement with both of those from Condon and Mitchell (1984) and Setia Gunawan et al. (2003).

## 3.4 Analysis of the COBRaS 21cm detections

### 3.4.1 Identified objects in CLASC

Working from the identifications made in Table 3.5, five of the objects detected within the COBRaS 21cm observations are well-studied massive stars or massive star systems. These include the LBV candidate (and possible multi-star system) Cyg OB2 #12, the O7.5III-I + O/B system A11, the quadruple stellar system of Cyg OB2 #5, and the well studied

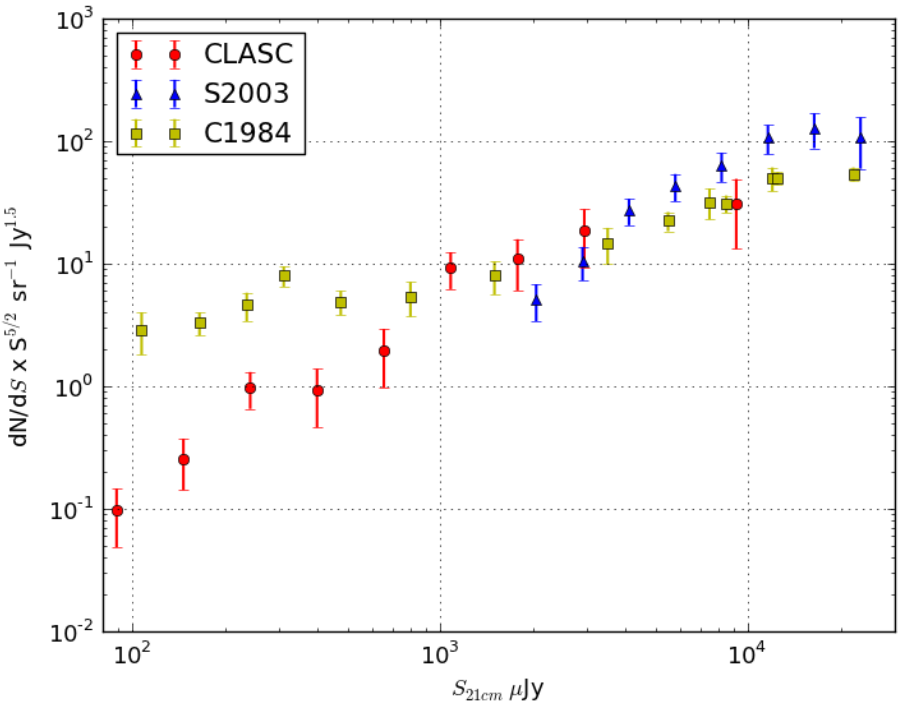



FIGURE 3.13: Normalised differential source counts. The data points corresponding to the sources within the CLASC are represented by red circles and are compared against the source counts from other surveys at 21cm by Setia Gunawan et al. (2003, blue triangles) and Condon and Mitchell (1984, yellow squares).

massive O star binaries of Cyg OB2 #8A and Cyg OB2 #9. A complete discussion of these objects and the analysis of their detected 21cm radio emission can be found elsewhere in this thesis. Cyg OB2 #12 is discussed at length within Chapter 4, whilst the other four known massive star multiple systems form the basis of much of the work presented in Chapter 5. The remaining identified objects within Table 3.5 are now discussed individually below.

#### 3.4.1.1 CLASC ID: #2

Previously detected at a wavelength of 86cm with a flux of 960 $\mu$Jy (Martí et al., 2007), this source is detected in the COBRaS 21cm Apr. 26th observations at a flux density of $213 \pm 38 \mu$Jy. The object has been classified as a very faint detection with a flux density of $101 \pm 32 \mu$Jy within the Apr. 11th observations due its low signal-noise-ratio (SNR = 2.7) and its small extent on the sky in comparison to the beam size (see Table C.1). Taking the 21cm flux density as measured on Apr. 26th and comparing it to $F_{86cm}$ as measured by Martí et al. (2007), the spectral index for this source is calculated to be $\alpha = -1.66$, suggesting a non-thermal radio spectrum. The object is marginally resolved within the COBRaS Apr. 26th observations with a slight elongation in the east-west direction. An extragalactic or Galactic origin for this source can not be determined from these COBRaS L-band observations alone.

#### 3.4.1.2 CLASC ID: #7

This relatively bright radio object has been detected at a 21cm flux density of $3.66 \pm 0.38$ mJy and $2.90 \pm 0.31$ mJy in the COBRaS Apr. 26th and 11th observations respectively. This does not suggest any large scale variability in the 21cm emission of this object over the 15 day period. The object has been previously detected at both 6cm and 21cm with the VLA at flux densities of $1.9 \pm 0.4$ mJy and $3.3 \pm 0.5$ mJy respectively (Martí et al., 2005), in good agreement with the 21cm flux densities found here. The authors were searching for possible 'hot spots' associated with the relativistic jets produced from the 'nearby' micro-quasar known as Cygnus X-3. They propose that the emission detected as source #7 in this work is indeed a hot spot candidate (HSC) and is associated with the

X-ray binary despite lying an angular distance of 7.07′ to the north. Martí et al. (2007) also detect a second HSC lying 4.36′ to the south of Cyg X-3 (and therefore not within the field of view covered by these COBRaS observations), which in combination with source #7 form an almost perfect alignment with the micro-quasar at a position angle of 1.8° ± 0.4°. This is in agreement with the position angle of the inner arcsecond radio jet (2.0° ± 0.4° as measured by Martí, Paredes, and Peracaula 2001). Moreover, the asymmetry observed between the north and south HSC is also similar to that observed in the inner jet components (Martí et al., 2005).

These observations of source #7 confirm the 21cm flux densities previously found for this HSC. The positional information derived here is also consistent with that found from the VLA observations taken in 2000 by Martí et al. (2005). Only a 0.144" offset is observed, which is within the ∼ 0.2" uncertainty associated with the VLA observations taken in 2000 (Martí et al., 2005). As mentioned in Section 2.4.3, Cyg X-3 is outside the primary beam of the Lovell antenna covered by COBRaS 21cm G pointing and does not feature in the wide field image of pointing G. As a result, its positional information cannot be determined, with which the proper motion of the HSC in relation to Cyg X-3 could have been investigated. With reference to Table C.1, Figures 3.10k and 3.10l, it is clear that source #7 appears to be fairly resolved within both epochs of the COBRaS L-band observations. However, the morphology represents nothing other than a spherically symmetric source. If this source is indeed associated with the micro-quasar Cyg X-3, then the detected 21cm emission (or at least part of the emission) must be a result of synchrotron radiation produced from the particles within the relativistic jets of a micro-quasar, that lies ∼ 14.8 parsecs (at a distance to Cyg X-3 of 7.2 kpc; Ling, Zhang, and Tang 2009) from its position. Whilst the symmetry with a second HSC to the south and their relation to the properties of the inner jets of Cyg X-3 suggest this is true, these 21cm COBRaS observations do not provide a conclusion as to the true nature of this object.

#### 3.4.1.3 CLASC ID: #9

This object is detected in the COBRaS 21cm Apr. 26th observations with a flux density of 520 ± 68 $\mu$Jy. Despite this, the object has failed to appear within the observations taken $\sim$ 15 days prior to that epoch, where a 3$\sigma$ upper limit on the flux density of $<$ 147 $\mu$Jy is placed on the object. This suggests that the 21cm emission of source #9 is able to vary significantly over short timescales. It has been identified to a previously detected source from Martí et al. (2007), but with a low likelihood ratio ($LR$) of just 0.8. Treating this identification to be true, the previous 86cm flux density of $\sim$ 1.45 mJy (Martí et al., 2007), in combination with the 21cm COBRaS flux density implies a negative spectral index, $\alpha = -1.13$. This, however, assumes that the object is not radio variable over the two epochs sampled by these observations. Additionally, source #9 has been matched to a source from the Chandra X-ray observations of Wright et al. (2014) with a value of $LR = 2.3$. The object therefore emits at both X-ray and radio wavelengths, its radio spectrum is likely non-thermal suggesting the presence of synchrotron radiation whilst its 21cm emission has been observed to vary over a period of $\sim$ 15 days. It is likely therefore that this object is of a Galactic origin. It is possible this object is a potential massive star binary system, however unlike the other four massive multiple systems observed within these COBRaS L-band observations it is not seen within the 2MASS survey.

#### 3.4.1.4 CLASC ID: #10

This object has been identified with source SBHW 81 from the previous radio observations by Setia Gunawan et al. (2003) with a very high $LR$. Source #10 (SBHW 81 hereafter) is marginally resolved in both of the COBRaS observation epochs with a slightly elongated structure to the south and to the east for the Apr. 26th and Apr. 11th observations respectively (see Figures 3.10m and 3.10n). With no detection at 86cm from the observations of Setia Gunawan et al. (2003), those authors classified SBHW 81 as type S/O (stellar/other) suggesting a Galactic origin. Furthermore they detected this object at 21cm with a flux density of 3.6 ± 0.5 mJy. The two 21cm fluxes here as measured with e-MERLIN are lower than this with values of 2.53 ± 0.27 and 1.55 ± 0.18 for Apr. 26th and Apr. 11th epochs

respectively. Despite the difference in sensitivity and likely $u, v$ coverage between these three observations, the variation between the flux densities is significant. A variation from 1.55 to 2.53 mJy in just 15 days argues against an extra-galactic origin for this source. Such a conclusion is less likely because the radio emission within a galaxy or AGN being sampled by these observations will often cover a huge physical distance (perhaps with the exception of compact AGN). This means an enormous amount of energy would be required to vary the produced radio emission over such large scales and thus one would expect the 21cm emission of an extra-galactic object to be relatively stable over short timescales.

#### 3.4.1.5 CLASC ID: #11

ID #11 has also been identified with a source detected from the radio observations of Setia Gunawan et al. (2003) known as SBHW 83. Detected in both observation epochs of the COBRaS L-band data as shown in Figures 3.10o and 3.10p, this source shows a huge amount of extended structure. The observations taken on Apr. 26th resolve three different components to the system. The 21cm flux density, $F_{21cm}^{26^{th}} = 3.59 \pm 0.42$ corresponds to the flux density of the main central component, i.e. source ID #11a (as measured with SEAC). Summing the measured flux densities of each of the detected components results in a total 21cm flux density of $5.76 \pm 0.66$ mJy. The Apr. 11th observations are only able to resolve the main component (ID # 11a) at a flux density of $900 \pm 147$ $\mu$Jy. The difference between the images from the two epochs is attributed to the different spatial scales sampled by their $u, v$ coverages (see Figures B.1c and B.1d for a comparison). Thus the resolved nature of the source and the under sampled inner baselines of the $u, v$ coverage means a large amount of the flux is resolved out of the Apr. 11th observations for SBHW 83. As a result the observations from either epoch may not make for a fair comparison.

The observations of Setia Gunawan et al. (2003) report flux densities at 21cm and 86cm of $8.4 \pm 0.5$ mJy and $62 \pm 3$ mJy respectively for SBHW 83. Whilst this 21cm flux density is comparable to that found from the Apr. 26th COBRaS observations (note that the object is unresolved in the observations from Setia Gunawan et al. 2003), the high 86cm flux density in combination with the 21cm flux density derived here results in a negative

spectral index, $\alpha = -1.6$. Note this is likely a lower limit for the spectral index since the COBRaS Apr. 26th observations are still likely missing flux on large spatial scales in comparison to the 86cm observations from Setia Gunawan et al. (2003). The non-thermal radio spectrum and extended morphological features suggest that SBHW 83 is most likely extra-Galactic in origin.

#### 3.4.1.6 CLASC ID: #17

Similarly to source ID#11 (or SBHW 83), source ID #17 has also been detected at both 21cm and 86cm from the observations of Setia Gunawan et al. (2003). They refer to this object as SBHW 90 and measure its 21cm and 86cm flux density to be 14.7 ± 0.8 mJy and 54 ± 4 mJy respectively. This 21cm flux is in excellent agreement with that derived from the COBRaS 21cm observations taken on Apr. 26th, $F_{21cm}^{26^{th}} = 14.66 \pm 1.49$ mJy. The 21cm flux density as measured from the Apr. 11th observations is slightly lower at 10.80 ± 1.10 mJy, however this again is likely due to the differences between the $u, v$ coverages between the two epochs. With reference to Figures B.1m and B.1n, it is clear that the Apr. 26th epoch samples a far greater extent of the $u, v$ plane and thus it is conceivable that a large amount of flux is missing within the Apr. 11th observations, when considering the heavily resolved nature of the source. SBHW 90 is the brightest source observed within these COBRaS L-band observations and the ratio between the source size ($Ss$) and the beam size is 32.0 and 8.6 for the Apr. 26th and Apr. 11th observations respectively, highlighting its resolved nature.

Setia Gunawan et al. (2003) determine a 21cm flux density of 19 mJy when convolved with the same beam size as their 86cm observations and hence derive a negative spectral index, $\alpha = -0.77$ for SBHW 90. The negative spectral index associated with this source is confirmed by using the 21cm flux density from the COBRaS Apr. 26th observations, giving a value of $\alpha = -0.94$. Similarly to ID#11 (SBHW 83), the heavily resolved nature of ID #17 (SBHW 90) and its negative spectral suggest an extra-Galactic origin for this source.

#### 3.4.1.7 CLASC ID: #18

Source ID #18 is another object detected across both of the COBRaS 21cm observation epochs with flux densities of 742 ± 89 $\mu$Jy and 438 ± 76 $\mu$Jy for those of the Apr. 26th and Apr. 11th respectively. With reference to Figures 3.10u and 3.10v, the object appears marginally resolved in both epochs with a $Ss$ to beam ratio of 3.29 and 1.43, again for the Apr. 26th and Apr. 11th observations respectively. This in combination with the large difference between $u, v$ coverage sampled by either observation epoch and the small difference between the flux densities imply that this source is not significantly variable over the 15 day period. ID #18 has only been previously detected within the X-ray observations of Wright et al. (2014) as a point source within their Chandra Legacy survey of the Cyg OB2 association. Whilst an extragalactic origin cannot be ruled out for this object, its inclusion within the X-ray point source catalogue from Wright et al. (2014) suggests a possible Galactic origin for this source.

#### 3.4.1.8 CLASC ID: #27

This object has also been previously detected at X-ray wavelengths from the observations of Wright and Drake (2009) and Wright et al. (2014). At detected 21cm flux densities of 299 ± 45 $\mu$Jy and 208 ± 44 $\mu$Jy from the Apr. 26th and Apr. 11th COBRaS observations respectively, this object doesn't appear to show any significant variation over the timescale of 15 days. Moreover the object appears marginally resolved in the Apr. 26th observations and slightly more compact in those of Apr. 11th. Again, its inclusion within the X-ray point source catalogues as observed with Chandra and the fact this object is not largely resolved lead to the suggestion of a possible Galactic origin for this object.

#### 3.4.1.9 CLASC ID: #33

ID #33 is another object previously identified as SBHW 109 from the observations of Setia Gunawan et al. (2003). It is only detected in the Apr. 26th epoch of the COBRaS observations and is found to contain two components, one to the south (109a) at a 21cm

flux density of 263 ± 50 $\mu$Jy and another to the north (109b) at a flux density of 118 ± 33 $\mu$Jy. The image of SBHW 109 from the Apr. 26th (Figure 3.12a) there includes a cross marking the position of SBHW 109 as given by Setia Gunawan et al. (2003). With a derived 21cm flux density of 3.0 ± 0.5 mJy from Setia Gunawan et al. (2003) it is clear that the COBRaS observations are resolving out the majority of the radio emission associated with the source. With no known detection of the source at other radio frequencies no spectral index information can be obtained. From a morphological perspective, this source resembles the radio emission from bi-polar jets associated with an Active Galactic Nuclei (AGN) and is therefore likely of an extragalactic origin.

#### 3.4.1.10 CLASC ID: #34

This object is another of those identified with the radio observations of Setia Gunawan et al. (2003). Otherwise referred to as SBHW 110, this source is detected in both epochs of the COBRaS 21cm observations (see Figures 3.10ao and 3.10ap). The 21cm flux density of SBHW 110 is shown to be consistent over this 15 day period with values of 3.92 ± 0.40 mJy and 3.52 ± 0.38 mJy for the Apr. 26th and Apr. 11th epochs respectively. Setia Gunawan et al. (2003) measure its 21cm flux density to be 2.2 ± 0.2 mJy but are unable to detect its emission at 86cm. The authors suggest the emission is likely thermal in nature and thus the source is of a Galactic origin. Within the COBRaS observations, SBHW 110 appears to be almost spherically symmetric and is fairly resolved with a $Ss$ to beam ratio of 5.0 and 4.0 for the Apr. 26th and Apr. 11th epochs respectively. There also appears to be some low level 21cm emission within the immediate vicinity of the source, however this is very close to the noise level. There currently exists no other detections of this source at any other frequencies within the literature. With no spectral index information and no significant variability over a short timescale (i.e. 15 days), it is difficult to determine if this object is associated with our Galaxy or not and further observations are clearly required to draw such a conclusion.

#### 3.4.1.11 CLASC ID: #35

ID #35 appears to be fairly resolved in both of the COBRaS observation epochs and shows a slight deviation from spherical symmetry due to elongation to the north west and north east within the Apr. 26th and Apr. 11th observations respectively. This source has been identified with SBHW 112, another object first detected within the radio observations of Setia Gunawan et al. (2003). Similarly to ID #10 (SBHW 81), the 21cm flux density of SBHW 112 is found to significantly vary over the ∼15 day period from a value of 1.42 ± 0.19 mJy (2014 Apr. 11th), to 2.95 ± 0.3 mJy (2014 Apr. 26th). Setia Gunawan et al. (2003) find a 21cm flux density of 2.2 ± 0.3 mJy but are unable to detect the source at 86cm. The authors suggest that the object is found within the galaxy and likely of a stellar nature. The short term variation in the 21cm flux density argues against this source being extra-Galactic in origin. Instead, the 21cm radio emission suggests that this object is situated within our Galaxy.

#### 3.4.1.12 CLASC ID: #38

This object has been identified as a source from the catalogue from Guarcello et al. (2013) with a $LR \approx 756$. This catalogue of proto-planetary disk stars within the Cyg OB2 association was compiled from a combination of observations from various IR surveys. This object in particular has been observed with the Spitzer space telescope at 3.6$\mu$m, 4.5$\mu$m, 5.8$\mu$m, 8.0$\mu$m and 24$\mu$m. Guarcello et al. (2013) determine an IR spectral index, $\alpha_{IR} = 0.547$ for the source and classify it as a class I YSO, which form ∼ 8.4% of their sample of 1843 disk stars. A given class of a YSO describes its evolutionary status. At these early evolutionary stages of class I YSOs, the central star of the object is surrounded by an accreting circumstellar disk and envelope which has yet to dissipate as occurs within the later evolutionary stages.

ID #38 is detected in the COBRaS 21cm observations taken on Apr. 26th with a flux density of 405 ± 58 $\mu$Jy. It appears fairly resolved with a $Ss$ to beam ratio of 2.69 and shows a slightly elongated morphology to the east (see Figure 3.10aw). The object is faintly detected in the Apr. 11th observations at a flux density of 177 ± 50 $\mu$Jy with a

signal-to-noise ratio 3.6. There is a slight difference between the $u, v$ coverage of pointing C between the two epochs (see Figures B.1e and B.1f) and as a result the variation in the flux density between the two epochs may not be significant. Further observations are required to derive its radio spectral index, from which a comparison can be made to that expected of a class I YSO.

#### 3.4.1.13 CLASC ID: #39

ID #39 has been detected in the COBRaS Apr. 26th observations with a 21cm flux density of 151 ± 38 $\mu$Jy. It is not detected in the Apr. 11th epoch due to the decreased sensitivity of these observations and the inherently low 21cm flux density associated with the source (and thus no variability can be found between the two epochs). The object is very compact and appears point like within the Apr. 26th observations with a $Ss$ to beam ratio of just 1.09. It has been identified with an X-ray source within the observations of Wright et al. (2014) and also to a 2MASS object (Cutri et al., 2003), both with a high likelihood ratio. It is likely that this source is therefore of a Galactic origin and possibly a member of the Cyg OB2 association.

#### 3.4.1.14 CLASC ID: #41

ID #41 has been identified with another source from the X-ray point source catalogue of Wright et al. (2014) with a $LR$ value of $\sim$ 540. It has been detected in the Apr. 26th epoch with a flux density of 1297 ± 158 $\mu$Jy and appears resolved with a $Ss$ over beam ratio of 5.75. With reference to Figure 3.12d, the object appears elongated in the south-west/north-east direction with a small amount of structure extending also to the north of the source. Interestingly, despite its relatively large measured flux density in the Apr. 26th observations, ID #41 is not detected at all within the Apr. 11th observations. From this epoch, a 3$\sigma$ upper limit of 225 $\mu$Jy is placed upon its flux density. The source is located towards the very edge of the wide field images of pointing C, explaining the large local RMS values of 38 $\mu$Jy and 75 $\mu$Jy for the Apr. 26th and Apr. 11th observations respectively. These COBRaS observations show that the 21cm emission from ID #41 is highly variable

over a 15 period which argues against a possible extragalactic origin for this object. Since this variation is so large over such a short period of time, this gives two possible scenarios for this source. It is either a short period massive star binary whose colliding wind region (CWR) gives rise to a large amount of non-thermal (synchrotron) emission (see Chapter 5) which is subsequently reduced (i.e. due to large amount of free-free absorption along the line of sight) a short time later at a different orbital phase, or this a flaring event due to some pre-main sequence type object. It is clear that further observations are required in order to decipher the physical properties of this object.

### 3.4.2 Un-identified objects in CLASC

Of the 41 objects detected within the COBRaS L-band observations, 22 were found without any previous identification. With measured flux densities at a single frequency, no spectral index information can be obtained for these objects. The Apr. 11th observations provide 11 of the unknown sources with a second flux density measurement. For the other half of the unknown sources, one must rely on a $3\sigma$ upper limit on the flux density to make a comparison between the two observation epochs. Due to the difference in the $u, v$ coverage between the two 21cm COBRaS observation epochs, the resulting images will be sensitive to slightly different spatial scales on the sky. Therefore, for those sources that are resolved, i.e. their source size to beam (in pixels) ratio, $Ss/\theta_{beam} \gtrsim 1.9$, this must be taken into consideration when comparing flux densities between the two observation epochs[2].

Table 3.7 shows the flux densities (or $3\sigma$ upper limit on the flux densities) and their $Ss$ to beam size ratio of the unknown sources across both observation epochs. The only possible classification of these sources must come from a qualitative assessment of their morphologies. Each source has been categorised according to the shape of the detected 21cm emission across both epochs where possible. Objects with a spherical or slightly elongated structure have been denoted (S/E) whilst those representing a bow shape like structure have been denoted with a (B). Sources in between either morphology have been given an (E/B) morphological type and finally those sources with extended structure are

[2]This definition of an 'extended' or resolved source has also been used in the analysis of the Westerbork radio survey by Setia Gunawan et al. (2003).

given an (Ex) type. A bow-shaped like structure is perhaps indicative of non-thermal (synchrotron) emission as a result of the collision between two massive star winds. This was found to be the case for Cyg OB2 #9, which is now known to harbour a CWR between its two massive star components (see Dougherty and Pittard 2006). Extended structure on the other hand is more likely to indicate a possible extra-galactic origin for a given source since large scale extended radio emission is readily observed in AGN and radio galaxies.

TABLE 3.7: CLASC sources without identification.

| ID | $F_{21cm}^{26^{th}}$ ($\mu$Jy) | $Ss_{26^{th}}/\theta_{beam}$ | $F_{21cm}^{11^{th}}$ ($\mu$Jy) | $Ss_{11^{th}}/\theta_{beam}$ | Morphology |
|---|---|---|---|---|---|
| 1 | 1574±166 | 4.86 | 1062±127 | 3.01 | S/E |
| 3 | 236±42 | 2.10 | 132±40 | 0.92 | B |
| 4 | 145±29 | 1.32 | 115±36 | 0.99 | B |
| 5 | 213±57 | 1.02 | <327 | - | E |
| 13 | 97±29 | 1.00 | <135 | - | E |
| 14 | 98±29 | 1.06 | <144 | - | B |
| 15 | 127±36 | 1.27 | <162 | - | B |
| 16 | 100±29 | 0.92 | <159 | - | B |
| 19 | 329±53 | 1.89 | <165 | - | E |
| 20 | 673±83 | 3.06 | 214±58 | 0.89 | E/B |
| 23 | 229±44 | 1.47 | 133±43 | 0.77 | E |
| 24a | 258±75 | 1.69 | <294 | - | B |
| 24b | 1444±188 | 6.93 | <294 | - | B |
| 25 | 245±44 | 1.61 | <150 | - | E/B |
| 26 | 2791±289 | 5.00 | 2281±247 | 3.32 | S/E |
| 28a | 327±69 | 1.95 | 302±81 | 1.38 | S/E |
| 28b | 1092±143 | 5.84 | <162 | - | Ex |
| 29 | 67±20 | 0.80 | <146 | - | E/B |
| 30 | 631±82 | 4.39 | 208±45 | 1.06 | E |
| 31 | 140±33 | 1.34 | <138 | - | B |
| 32 | 1890±214 | 14.44 | 483±98 | 2.12 | Ex |
| 36 | 1003±120 | 3.83 | 656±101 | 1.94 | S/E |
| 37 | 1472±174 | 4.98 | 881±155 | 1.90 | E/B |
| 40 | 1627±184 | 8.65 | <171 | - | Ex |

Morphologies: (S/E) spherical/elongated; (B) bow shaped; (E) elongated; (Ex) extended.

### 3.4.3 Discussion

The COBRaS L-band observations taken on 2014 Apr. 26th have shown that 41 objects within a $\sim$ 0.48 deg$^2$ area of the centre of the Cyg OB2 association emit 21cm radio

emission above a flux density limit of $\sim$ 70 $\mu$Jy[3]. Additional observations, accounting for $\sim$ 1/3 of the on-source time of the Apr. 26th epoch were made $\sim$ 15 days previously and detected 25 of these objects. Moreover, the majority of the sources in the Apr. 26th observations are detected at higher flux densities than in the Apr. 11th observations (with the exception of ID #22, a radio variable colliding wind binary system otherwise known as Cyg OB2 #8A). Unfortunately, there exists no detected source whose radio flux density can be reliably deemed to be constant between the two observation epochs. As a result it is difficult to test the reliability concerning the observed variation of a number of sources within CLASC.

The only characteristics that can conceivably account for flux density variations between the two epochs include the difference between the $u, v$ coverage and sensitivity of the data sets. Appendix B, shows the $u, v$ coverage of each pointing for either epoch. For a number of pointings there is a significant difference in the regions in the $u, v$ plane sampled by either epoch. For compact, unresolved sources, this is not generally considered as a problem, however the majority of sources detected in CLASC are resolved and contain extended structure. In these cases, an under-sampled $u, v$ coverage will lead to images that are sensitive to flux on fewer spatial scales resulting in a net lowering of the integrated source flux density. The sensitivity of the images comes into play when considering the pixel-by-pixel flux extraction method implemented by SEAC. For low signal-to-noise sources, the true extent of the source can often be hidden within the noise of the image, as a result their flux is underestimated. This however, is only the case for low signal-to-noise sources (i.e. SNR $\leq$ 7; Peck 2014).

To give a possible constraint on the observed variation between either epoch as a function of the $u, v$ coverage, consider a 'worst case scenario'. ID #17 (see Figure 3.10s), previously identified as SBHW 90 by Setia Gunawan et al. (2003), is the brightest object detected within the CLASC. This can be considered as the best 'worst case' scenario, since it is heavily resolved (containing a lot of extended emission) and there is a large difference in the $u, v$ coverages between the Apr. 11th and Apr. 26th epochs of pointing G in which it

[3]Note that this flux density limit applies to the central regions of each pointing and increases towards their outer regions.

was detected. Furthermore, it was detected in Setia Gunawan et al. (2003) with a 21cm flux density of $14.7 \pm 0.8$ mJy and in the Apr. 26th COBRaS data with a 21cm flux density of $14.66 \pm 1.49$ mJy, i.e. if neglecting the Apr. 11th flux density (which we assume to be affected by an under-sampled $u, v$ plane), the source appears to be constant over the $\sim$ 18 years between these observations. If one assumes that this source is therefore also non-varying over a 15 day period, the difference in $u, v$ coverages between the two COBRaS 21cm epochs, accounts for a variation in the source flux by a factor of $\sim 1.36$ in the 'worst case' scenario. In reference in Table 3.3, if ignoring the low signal-to-noise objects (i.e. those with a SNR $\leq 7$), this still means that 21 of the 41 detected sources have been found to be variable over a 15 period when accounting for this 'worst case' scenario in regards to the variation caused by $u, v$ coverage differences. This is over 50% of the observed population and a result which calls for a large number of follow up observations for these sources in order to further monitor their radio variability.

In regards to radio variability of objects within CLASC, there exist five objects which stand-out over the rest. These include object IDs #24, #28, #32, #40 and #41. All of which are previously unidentified sources except for ID #41 which has been detected also in X-rays. For each of these objects, their 21cm flux density between the two COBRaS observation epochs differ by $\gtrsim 4$. In fact, only object ID #32 was detected in the Apr. 11th epoch and for the remaining four objects only upper limits on their 21cm flux densities could be obtained. Additionally these objects are all heavily resolved and showcase large amounts of extended emission at 21cm, similar to the case of ID #17 (i.e. SBHW90). Whilst flaring, pre-main sequence stars are expected to show radio variations on these scales, the probability of observing five of these objects in just one observation epoch is unlikely. Short period, PACWB systems could also potentially give rise to radio variability of these timescales, however they are generally seen to be compact like objects, as evidenced by the four known PACWB systems detected in these e-MERLIN observations. The similarity in the extended morphology and rapid 21cm flux density variations of these objects is suggestive that they are all of a similar nature. Follow up observations are vital in order to decipher the puzzling physical attributes observed in these five systems.

## 3.5 Conclusions and future work

The COBRaS 21cm observations were acquired throughout the spring of 2014 over two epochs separated by $\sim$ 15 days. The observations consist of seven over-lapping pointings covering an area of the sky of $\sim 0.48^{\circ} \times 0.48^{\circ}$, centred on the heart of the Cyg OB2 association. The source extraction process was applied to each of the seven wide field images and was carried out by a purpose built Source Extraction Algorithm for COBRaS (SEAC) that utilises a floodfill algorithm to detect and measure the positions and integrated flux densities of any 'island' above the set threshold levels, $\sigma_s$ and $\sigma_f$. Searching down to a floodfill threshold level, $\sigma_f = 3.0$ (and in some cases $\sigma_f = 2.6$ in the subsequent targeted images of individual sources), a total of 41 individual sources were detected within the 21cm COBRaS observations. 25 of these sources were detected within both epochs and a total of 19 were matched to previously identified sources from various surveys, catalogues and observations from within the literature.

Five of the identified sources include the massive star (and LBV candidate) Cyg OB2 #12, which is discussed at length within Chapter 4, and the four massive star binary systems Cyg OB2 #5, Cyg OB2 #8A, Cyg OB2 #9 and A11, which form the basis of the work presented in Chapter 5. No known single massive stars (with the possible exception of Cyg OB2 #12), have been detected within these COBRaS 21cm observations (see Chapter 4). Evidently, these observations are sensitive to the non-thermal emission from colliding wind binary systems. Each of the fourteen remaining identified objects (excluding the five massive star objects) have been discussed individually and suggestions to their origin (i.e. either Galactic or extra-Galactic) have been made (where possible) by considering their flux densities between the two epochs, the extent to which they are resolved, their morphology and their previous identification(s). Of these 14 objects:

(i) Two (ID #2 and #7) are of an unknown origin based upon these COBRaS observations alone. ID #7 however has been previously suggested as a HSC associated with the relativistic jets of the micro-quasar Cyg X-3.

(ii) Three (ID #11, #17 and #33) are suggested to be of an extragalactic origin, perhaps

associated with AGN. This hypothesis was based on their extended morphology and a lack of any significant variability between the two COBRaS 21cm observation epochs.

(iii) Nine (ID #9, #10, ID #18, #27, ID #34, #35, ID #38, #39 and #41) are likely situated within our Galaxy since they have either been matched (to a confident likelihood ratio) to at least one previous source also thought to be Galactic., or they have been found to vary significantly in their 21cm flux densities between the two COBRaS observations epochs. Additionally, of these nine objects:

(a) ID #38 is a confirmed class I YSO from the work of Guarcello et al. (2013).

(b) ID #41 shows an extreme variation (an increase by a factor $> 5.76$) in its 21cm flux over the 15 day period between epochs reminiscent of a possible flaring event in a pre-main sequence type object or a potential small period massive star binary system imaged at significantly different orbital phases.

(c) The remaining seven objects (ID #9, #10, ID #18, #27, ID #34, #35 and ID #39) are consistent with the non-thermal 21cm emission observed from massive star binaries or YSOs.

(d) Only one (ID #39) was matched to a source from the 2MASS survey, however its position on the J-H vs H-K plot as shown in Figure 3.6 does not lie within the region covered by the OB candidate stars.

In addition to those detections with previous identifications, 21 sources are newly discovered objects that come without any identification. A qualitative assessment of their morphologies has been made and is presented in Table 3.7. However without further information, a limited number of conclusions can be drawn for these objects.

The COBRaS C-band observations scheduled to commence in the beginning of 2017 will provide 6cm flux densities down to $1\sigma$ sensitivities of $\sim 3\mu$Jy. These data will provide spectral index information for the 41 sources detected within the L-band observations which will lead to a more quantitative assessment of the nature of each object. In addition to these C-band observations, further information can be obtained from the 21cm observations presented here. By an assessment of the $u, v$ coverage between the two observation epochs

for each of the seven pointings (see Figures B.1a - Figures B.1n) an aptly chosen Gaussian taper can be applied to the images. In choosing a $u, v$ taper that best represents both epochs, each source detected within both observation epochs can be re-imaged. This may provide a fairer comparison between the resulting flux densities of each source over the two epochs. Thus obtaining a better handle on their variability over the 15 day period, from which various conclusions can be drawn regarding their origin and possible nature.

For those sources detected with an X-ray counterpart from the Chandra observations of Wright, Drew, and Mohr-Smith (2015), a comparison can be made between their X-ray and radio luminosities. A large number of these objects have been branded as potential YSOs or potential massive star binary systems. Thus, a comparison between their 21cm flux densities and X-ray luminosities may able to differentiate between these two possibilities by comparing the two values to that predicted from the Gudel-Benz relation (Benz and Güdel, 1994) for active stars. The eight potential YSOs or massive star binaries are discussed further in Chapter 5 of this thesis. It is clear that both the unknown and previously identified sources will benefit from follow up observations in order to better monitor their radio variability, which will bring further evidence regarding the nature of these objects.

# Chapter 4

# 21cm (L-band) observations of single massive stars in Cyg OB2

In astrophysics, our understanding of massive stars is fundamental. The effect of their formation, evolution and subsequent destruction has been crucial in shaping the universe to what it is today and undoubtedly what it will become in the future. They are responsible for the re-ionisation of the universe, affect the structure and evolution of the interstellar medium (ISM), play a significant role in galaxy formation whilst also providing the raw materials for the next generation of stars and planetary systems. A prominent feature of massive stars is their powerful stellar winds which shed huge amounts of mass per unit time. Their mass-loss rate, $\dot{M}$, a crucial parameter for stellar, galactic and cosmic evolution models, dictates their own evolution whilst at later evolutionary stages (post-main sequence) episodic mass-loss events can drastically alter their eventual end point. This leads to the production of core collapse supernovae, neutron stars and even black holes. Our knowledge of stellar mass-loss still remains one of the most uncertain parameters within massive star evolution and it must be understood quantitatively if it is to be included within broader astrophysical applications.

The majority of the work contributing to this chapter is presented in a recent paper, published in the Monthly Notices of the Royal Society (MNRAS), titled 'e-MERLIN 21cm constraints on the mass-loss rates of OB stars in Cyg OB2' (Morford et al., 2016). As the $1^{st}$ author, I wrote the entirety of the paper which was subsequently reviewed and iterated upon by each of the co-authors. Moreover, the work done in the calibration and imaging

of the data was carried out by myself with the aid of Danielle Fenech. Section 4.3 and the following discussion and conclusions (Section 4.4) are drawn almost directly from the paper by Morford et al. (2016).

This chapter first begins by summarising much of the background theory stated in Section 1.2, regarding the winds of massive stars, including their driving mechanism, typical characteristics and key observational diagnostics. In also summarising the current uncertainties behind stellar mass-loss, the chapter brings the focus onto the radio emission from single massive stars in Cyg OB2, including the tools required to estimate $\dot{M}$ from such emission (Section 4.2). Section 4.3 then compares the results gained from the COBRaS L-band observations to both those predicted by current theoretical work and other observational methods, from which constraints are placed upon any wind structure that may exist within the outflows of single massive stars.

## 4.1 The winds of hot stars: a summary

Massive stars have a stellar wind that can be described as a continuous flow of material that is blown radially outwards from the photosphere and into the surrounding ISM. Their mass-loss rate describes the amount of mass per unit time, which escapes the star via its stellar wind. The outflow is driven by a mechanism known as radiation line driving (Lucy and Solomon, 1970; Castor, Abbott, and Klein, 1975), whereby momentum is transferred (via scattering and absorption processes) from the stellar radiation field (i.e. the photons) to metal ions in the outer photosphere. The metal ions impart (or share) this momentum with the more abundant species of H and He resulting in a net outwards acceleration of the wind material (see Section 1.2.1 for further details).

A massive star's stellar wind gives rise to numerous observational signatures, which can be used to infer the physical properties of the wind such as the mass-loss rate and in some cases the terminal velocity of the wind. These include UV resonance lines, whose line profiles can be used to infer the column density of the wind and furthermore the mass-loss rate (Lamers, Cerruti-Sola, and Perinotto, 1987); the modelling of optical and near IR spectral

lines, whose properties depend on the square of the density of the emitting material (e.g. see Herrero et al. 2001; Herrero, Najarro, and Puls 2003); observations of the IR, mm and radio continuum whereby the free-free thermal emission also depends on the square of the density (Wright and Barlow, 1975); and X-ray observations where the attenuation in their line profiles has been linked to the wind column density (Cohen et al., 2010).

Each diagnostic probes a different region of the wind and a discordance has often been found between their inferred mass-loss rates (Drew, 1990; Fullerton, Massa, and Prinja, 2006; Puls et al., 2006; Muijres et al., 2011). This has led to the current uncertainty around the mass-loss rates of massive stars and has further led to the introduction of wind structures into theoretical models of massive star winds. Such models include 'clumps' of over dense material which can lead to an overestimate of the stellar mass-loss rate from those diagnostics which depend strongly on the density (i.e. $\rho^2$ diagnostics such as the optical and near IR line profile modelling and the IR, mm and radio continuum emission; see e.g. Puls, Vink, and Najarro 2008a). The uncertainty surrounding massive star wind structures and their effect on the mass-loss rate of these objects is much of the motivation behind the COBRaS project (see Section 1.6) and much of the work presented in this thesis.

### 4.1.1 The importance of stellar mass-loss and its current uncertainty

Two key areas of doubt in regard to the winds of massive stars still remain. Firstly, the weak wind problem that encompasses late O-dwarfs stars with unexpectedly small mass-loss rates in comparison to those predicted from the $\dot{M} \propto L^X$ power-law relations from bright O-stars (see e.g. Kudritzki et al. 1991; Puls et al. 1996). It was found by Martins et al. (2005) that including X-ray emission in low density stellar wind models leads to a degeneracy in the derived mass-loss rate from UV lines. It has been further suggested that the NIR Br$\alpha$ line may provide a mass-loss estimate that is free from the effects of X-ray emission, advection and adiabatic cooling in the case of thin stellar winds. Further work must be done to discover the exact parameters that differentiate stars with weak winds from those with 'normal' winds, at least in the case of O-type dwarf stars.

Secondly, despite many investigations into the different structures found within the winds of hot stars and their effect on various mass-loss diagnostics, wind clumping and therefore mass-loss estimates are still a major uncertainty within hot stars. The radial stratification of the clumping factor is of particular concern, with considerable observational evidence suggesting a large amount of clumping near to the stellar surface. The theoretical predictions of Runacres and Owocki (2002) indeed find that LDIs lead to the onset of clumping near to the stellar surface which can then survive to substantial distances ($R \sim 100R_\star$). They predict that beyond $\sim 20 - 30R_\star$, the line driving force has little effect on maintaining such structures and their evolution is dictated by the gas dynamics, i.e. the battle between the outwards pressure force, expanding individual clumps and the supersonic collision between different clumps acting to further compress them. Considerable effort must be made to tie together such theoretical predictions with observational evidence. From an observational perspective measurements of the clumping factor and the porosity length must be obtained across the electromagnetic spectrum (and therefore at different regions of the wind) for a whole host of spectral types.

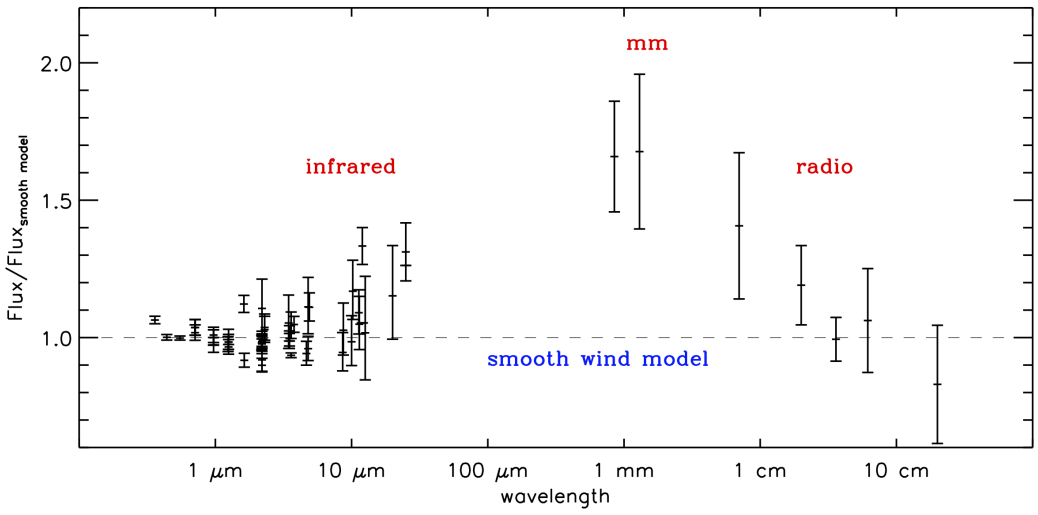



FIGURE 4.1: Fluxes obtained for the B0 Ia star $\epsilon$ Ori from observations in optical to the radio, normalised to those expected from a smooth wind ($f_{cl} = 1$). The plot essentially represents how the clumping factor changes at different wavelengths (i.e. radially outwards). Image taken from Blomme (2011).

It is clear that further observational evidence is required to measure the mass-loss rates

of massive stars at different regions in the wind. In this way, a picture of the radial stratification of the clumping factor can be built, as has been done with the nearby B0 Ia star $\epsilon$ Ori (see Figure 4.1). The plot shows how the flux, in comparison to that assuming a smooth wind (i.e. $f_{cl} = 0$) increases from the optical to longer wavelengths, peaking at mm wavelengths before decreasing towards longer radio wavelengths. It is important that such constraints are created for each massive star spectral type, such that a radial dependance of the clumping factor can be characterised as a function of stellar parameters, allowing for reliable determinations of stellar mass-loss. The diagram shown in Figure 4.2 highlights the regions in the wind that are dominated by emission at a given waveband, and hence shows the particular wind regions sampled when diagnosing a mass-loss rate from observations at a particular wavelength.

If wind clumping is assumed to contribute to the observed fluxes in massive stars (in the case of $\rho^2$ diagnostics), then the consequent downward revision of these empirical mass-loss rates would have a profound effect on current models of stellar evolution. The consequence however, depends on the extent to which the empirical mass-loss rate may have to be altered and hence the extent of wind clumping. Since $\dot{M} \propto \sqrt{f_{cl}}$, then clumping factors of $\sim 4$ would imply a mass-loss reduction of only $\sim 2$. For O- and early B-stars at least, this is still in agreement with those currently predicted by theory (see e.g. Vink, Koter, and Lamers 2000; Vink, Koter, and Lamers 2001). Clumping factors larger than $\sim 10$ however, will have a more profound effect on $\dot{M}$ values, undermining those expected from theory and having serious consequences for stellar evolution models and their greater application upon astrophysics.

It has been shown that a relationship exists between the mass-loss rate of OB stars and the content of metals within their atmospheres. From observations, Mokiem et al. (2007) derived a power law dependance between the mass-loss rate and metallicity, $\dot{M} \propto Z^m$. If adopting the theoretical relation between the wind terminal velocity and metal content, $v_\infty \propto Z^{0.13}$ as defined by Leitherer, Robert, and Drissen (1992), they found that $m = 0.83 \pm 0.16$ for 'smooth' (un-clumped) stellar winds. This is somewhat intuitive in regards to that predicted from radiative line driven wind theory, since metal ions provide the main driving mechanism in stellar mass-loss. This result however, implies that the mass-loss

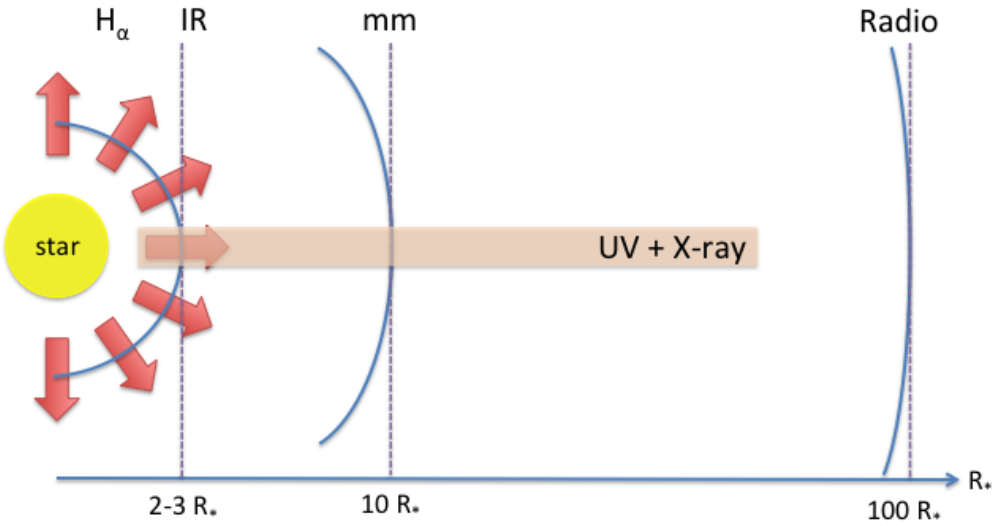



FIGURE 4.2: A schematic diagram to illustrate the regions in the wind which are dominated by emission at a given waveband from which one can diagnose the mass-loss rate of the wind.

rates of early type stars are increasing as a function of metal content and thus with ever increasing age of the universe. If a subsequent downward revision of mass-loss rates is necessary, this relation between $\dot{M}$ and metallicity could implicate serious consequences for the early universe and the epoch of re-ionisation, for which massive stars have been thought to be responsible for (Haehnelt et al., 2001; Wyithe and Loeb, 2003).

## 4.2 Radio emission from massive stellar winds

### 4.2.1 The winds of massive stars

As previously mentioned in Section 1.2, free-free radiation at radio wavelengths is emitted from winds of massive stars. Stellar winds are comprised of a hot plasma, the bulk majority of which takes the form of ionised hydrogen. An abundance of charged particles, i.e. free protons and free electrons, facilitates the production of free-free thermal emission across a broad frequency range. This has been previously referred to as the IR, mm and radio continuum and is a direct result of 'Bremsstrahlung' (literally meaning 'braking radiation') due to the interaction between two charged particles.

A negatively charged electron, of initial energy $E_i$, passing within close proximity of a positively charged proton, will change direction due to the electromagnetic force shared between two particles of opposing charge. This change of trajectory will cause the electron to slow down, resulting in a net reduction in its kinetic energy. To maintain a conservation of energy, the interaction causes the emission of a photon of energy $h\nu = E_i - E_f$, at some frequency $\nu$ (where $E_f$ is the final energy of the electron and $h$ is the Planck constant). It is this mechanism that gives rise to the free-free emission observed in the winds of hot stars and ranges from IR wavelengths through to radio. The process is also termed as 'thermal' as it depends on the thermal energy of the electrons and nuclei, which are assumed to be in thermal equilibrium with one another. The mean thermal energy of the electrons in a plasma of temperature, $T$, can therefore be written as:

$$E_e = \frac{3kT}{2} \tag{4.1}$$

where $k$ is the Boltzmann constant. For a typical O-type star wind of temperature $T \sim 4 \times 10^4$, $E_e \sim 5$ electron volts (eV). This is significantly more than the energy of a photon emitted at a typical radio frequency ($\nu \sim 5$ GHz), $E_\gamma = h\nu \sim 2 \times 10^{-5}$ eV. The thermal dependance of this process implies that the produced emission spectrum increases with increasing frequency, i.e. thermal free-free emission produces a positive spectral index, $\alpha$ (in the positive sign convention).

### 4.2.2 The thermal free-free emission spectrum from early-type stars

In terms of utilising the observed radio flux from the winds of massive stars to derive their mass-loss rate, one must know the relationship between the free-free thermal emission and the stellar wind properties. The derivation of the thermal free-free emission spectrum from early type stars and its dependance upon the stellar and wind parameters was found independently by Wright and Barlow (1975), Panagia and Felli (1975) and Olnon (1975). Prior to their seminal work, it was long expected that the IR, mm and radio continua from massive stars was a result of thermal free-free emission. Observations of many of these objects showed a positive slope in the log flux density versus log frequency plot (i.e.

a positive spectral index), which led to much of the motivation behind their work. A concise, simplified version of this derivation is given below, however for further details and a thorough derivation, please refer to the three original manuscripts.

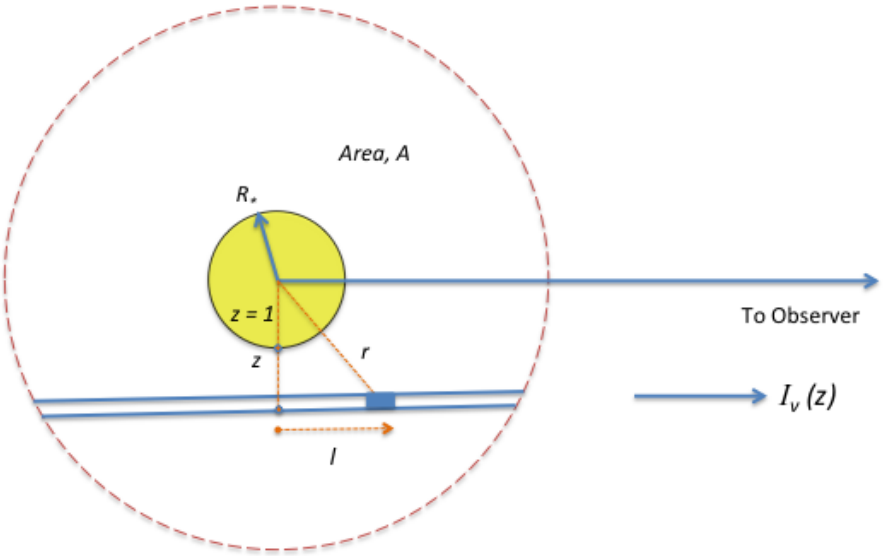



FIGURE 4.3: A schematic illustration to highlight the geometry of the model used to derive the free-free thermal emission spectrum throughout the winds of early-type stars.

The wind is assumed to be of a 'uniform flow', i.e. it is comprised of an ionised gas of number density $n_e$, that is ejected continuously outwards from a radius $R_\star$ and at constant velocity $v_\infty$. The equation of mass continuity applies (Equation 1.1) and the wind maintains a constant temperature, $T$ throughout. A simple illustration of the geometry used in the model is shown in Figure 4.3. The intensity of the emergent radiation from the thin cylindrical shell (as shown in Figure 4.3), under LTE can be written in terms of the planck function, $B_\nu(T)$:

$$I_\nu(T) = \int_0^{\tau_{max}(z)} B_\nu(T) e^{-\tau} d\tau = B_\nu(T)(1 - e^{-\tau_{max}(z)}) \tag{4.2}$$

The flux density from that emergent radiation received by the observer can be written as:

$$dF_\nu = \frac{I_\nu(T)}{D^2} 2\pi z = \frac{B_\nu(T)}{D^2}(1 - e^{-\tau_{max}(z)}) 2\pi z \tag{4.3}$$

where D is the distance between the object and the observer. The optical depth along the cylinder is given by:

$$\tau_\nu(z) = \kappa_\nu(T) \int_{-\infty}^{+\infty} n^2 dl = \frac{\kappa_\nu(T)\gamma\pi A^2}{2z^3} \qquad \text{for z} \geq 1 \tag{4.4}$$

where $n$ is the number density of gas in the wind, A is the area encompassed by the gas, i.e. $A = nr^2$ and $\gamma$ is the ratio of electron to ion density. The linear free-free absorption coefficient $\kappa_\nu(T)$ is that given (in *cgs* units) by Allen (1973):

$$\kappa_\nu(T) = 3.7 \times 10^8 \left(1 - e^{\frac{-h\nu}{kT}}\right) \overline{Z^2} g_\nu(T) T^{-1/2} \nu^{-3} \tag{4.5}$$

where $\overline{Z^2}$ is the mean ionic charge and $g_\nu(T)$ is the Gaunt factor. The flux received by the observer across the entire stellar wind is obtained by integrating Equation 4.3. The solution is vastly simplified under the assumption of large mass-loss rates and at low frequencies (please refer to Wright and Barlow (1975) for the full solution) and is written as:

$$F_\nu = 2.66\pi \left(\frac{\pi}{2}\right)^{2/3} \gamma^{2/3} \frac{A^{4/3}}{D^2} B_\nu(T) K_\nu^{2/3}(T) \tag{4.6}$$

This can be further simplified by first expressing the area of the emission region, $A$ in terms of the mass-loss rate by using the equation of mass continuity (Equation 1.1):

$$A = nr^2 = \frac{\dot{M}}{4\pi\mu m_H v_\infty} \tag{4.7}$$

where the density inside the shell of area $A$ has been expressed as $\rho = n\mu m_H$, i.e. in terms of the mean molecular weight, $\mu$, the number density of the gas $n$ and the mass of hydrogen, $m_H$. Inserting the expression for the free-free absorption coefficient then using this expression and further taking the planck function, $B_\nu(T) = 2\nu^2 kT/c^2$ in the

Rayleigh-Jeans limit ($h\nu << kT$), Equation 4.6 reduces to:

$$F_\nu = 23.2 \left( \frac{\dot{M}}{\mu v_\infty} \right)^{4/3} \frac{1}{D^2} \left( \gamma g_\nu(T) \nu \overline{Z^2} \right)^{2/3} \tag{4.8}$$

which is expressed in units of Janskys (Jy; 1 Jy= $10^{-26}$ W m$^{-2}$ Hz$^{-1}$), $\dot{M}$ is in M$_\odot$yr$^{-1}$, $D$ in kpc, $v_\infty$ in kms$^{-1}$ and $\nu$ in Hz. This equation shows that the flux received by an observer due to the thermal free-free emission is dependant on $\nu^{2/3}$. However, the Gaunt factor, $g_\nu$ is weakly dependant on the frequency of the emission, and typically reduces the spectral index, $\alpha$ of the free-free emission to $F_\nu \propto \nu^\alpha$, where $\alpha = 0.6$. This is in good agreement to IR, mm and radio continuum observations of the stellar winds of early type stars undergoing mass-loss (Wright and Barlow, 1975).

### 4.2.3 The effect of clumping on the thermal free-free emission spectrum

In Section 1.2.2, it was revealed that particular structures such as small-scale, optically thin clumps, are thought to exist within the winds of massive stars. It was shown, by the introduction of the clumping factor, $f_{cl}$ that the amount of wind clumping can be quantified. Moreover, for a $\rho^2$ diagnostic (such as a measurement of the free-free thermal emission), the mass-loss rate, $\dot{M}$ scales with the square root of the clumping factor.

For massive stars, Equation 4.8 relates the observed flux from thermal free-free emission to the mass-loss rate. This assumes a smooth, homogeneous stellar wind, free from any density structures. Equation 4.8 must therefore be modified to include $f_{cl}$, in order for observations of the IR, mm and radio continuum in early type stars to account for any wind clumping. This is achieved by inserting the dependance of the mass-loss rate on $f_{cl}$, such that Equation 4.8 becomes:

$$F_\nu = 23.2 \left( \frac{\dot{M}\sqrt{f_{cl}}}{\mu v_\infty} \right)^{4/3} \frac{1}{D^2} \left( \gamma g_\nu(T) \nu \overline{Z^2} \right)^{2/3} \tag{4.9}$$

In the case of a smooth wind, $f_{cl}$ takes a value of unity, and the original form of Equation 4.9 as given in Wright and Barlow (1975) is recovered. For any value of $f_{cl} > 1$, over-dense clumps are present within the wind, which has the effect of increasing the observed thermal free-free flux, for a given mass-loss rate. Hence, in utilising the thermal free-free radio emission to diagnose the mass-loss rate of a hot star, a given radio flux will either imply a maximum $\dot{M}$ value with no clumping (i.e. $f_{cl} = 1$), or a lowered value of $\dot{M}$ with $f_{cl} > 1$.

### 4.2.4 Predicting the radio flux from massive star winds

In preparation for using the COBRaS 21cm observations to derive the mass-loss rates of single massive stars in the Cygnus OB2 association, it is first useful to make a prediction of the L-band, 21cm radio emission for each single massive star within the field of view of the observations. These predictions can be used to determine the likelihood of detecting the 21cm emission from such objects and furthermore they offer a direct comparison to the observed 21cm flux. The theoretical prediction of the 21cm flux observed from the winds of single massive stars can be made using Equation 4.8. This requires prior knowledge of the stellar mass-loss rate, terminal velocity of the wind, the distance to the object and various parameters that describe wind ionisation conditions.

A prediction of the mass-loss rate for a particular early-type star can be made using the relations calculated from a theoretical grid of models by Vink, Koter, and Lamers (2000) and Vink, Koter, and Lamers (2001). The mass-loss recipe is a function of the stellar luminosity, mass, effective temperature, metallicity and the ratio of the terminal velocity to the escape velocity. These include two recipes, for stars either side of the so called 'bi-stability jump', which occurs at a $T_{eff} \sim 25000$ K (for the precise formulation of these relations please refer to Equations 24 and 25 in Vink, Koter, and Lamers 2001). This 'jump' is observed as a steep decrease in the terminal velocity of the stellar wind from a value of $v_{\infty} \simeq 2.6 v_{esc}$ to $v_{\infty} \simeq 1.3 v_{esc}$, which is also accompanied by an increase in the mass-loss rate (see Vink, Koter, and Lamers 1999 for more details).

Computing the mass-loss rate of O- and B-type stars from the relations given in Vink, Koter, and Lamers (2001) can be achieved using a theoretical approach that relies on prior knowledge of the object's spectral classification. The spectral types (and luminosity class) of all the known OB stars within the field of view of the COBRaS L-band observations has been amalgamated into the Cyg OB2 OB Star Catalogue (see Section 3.2.1). From these spectral classifications, one can derive the various stellar parameters used within the Vink, Koter, and Lamers (2001) recipe from calibrations.

For the O-stars, the stellar parameters ($T_{eff}, L_\star, M_\star, R_\star$, log$g$) are obtained through the calibrations of Martins, Schaerer, and Hillier (2005), who took a nLTE approach assuming solar metallicity, taking both wind effects and line-blanketing into account. Similarly for the B supergiant stars, the same stellar parameters are derived from the calibrations of Searle et al. (2008). Values of the wind terminal velocity as a function of spectral classification were obtained from Prinja, Barlow, and Howarth (1990) for both the O and B stars and the escape velocity, $v_{esc}$ could be calculated using the relation:

$$v_{esc} = \sqrt{\frac{2GM_{eff}}{R_\star}} \tag{4.10}$$

where the effective mass $M_{eff} = M_\star(1 - \Gamma_e)$, is defined by the Eddington parameter, $\Gamma_e$:

$$\Gamma_e = 7.66 \times 10^{-5} \sigma_e \left(\frac{L}{L_\odot}\right)\left(\frac{M_\star}{M_\odot}\right)^{-1} \tag{4.11}$$

which in turn is a function of the electron scattering cross-section, $\sigma_e = 0.34, 0.32$ and $0.31$ cm$^2$g$^{-1}$ for stars of $T_{eff} \geq 35$, $30 \leq T_{eff} < 30$ and $T_{eff} < 30$ kK respectively (Lamers and Leitherer, 1993).

Having assembled the necessary parameters, the mass-loss recipe as given by Vink, Koter, and Lamers (2001) can be used to calculate $\dot{M}$ for O-type dwarfs, giants and supergiants, and B-type supergiant stars. The top panel in Figure 4.4 shows a plot of theoretically calculated $\dot{M}$ values for each of these spectral types, as calculated with the Vink recipe using an updated solar metallicity, $Z = 0.013$ (as given in Asplund et al. 2009). It it

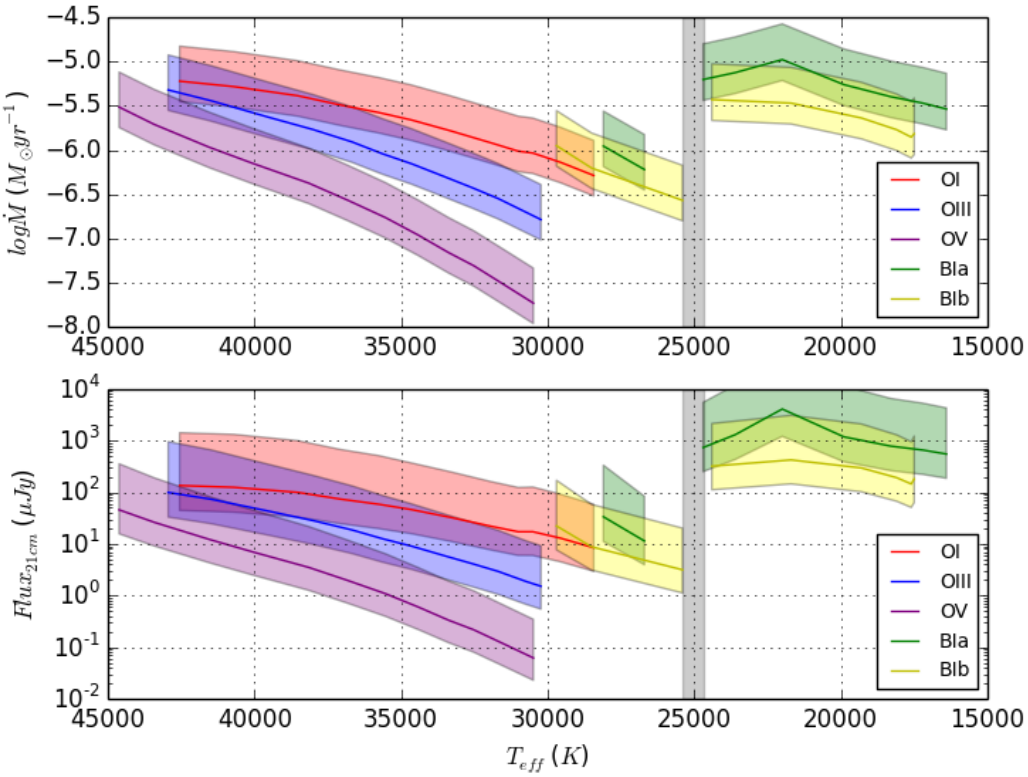



FIGURE 4.4: Top - Theoretically calculated values of $\dot{M}$ using the prescription from Vink, Koter, and Lamers (2001), plotted as a function of effective temperature for O-type dwarfs, giants, supergiants and B-type supergiant stars. Bottom - the predicted 21cm flux assuming a smooth wind model and the theoretical $\dot{M}$ values as a function of effective temperature (and spectral classification). The grey region around 25 kK represents the approximate region of the observed bi-stability jump in early type stars, whilst the shaded coloured regions represents the uncertainty in the respective $\dot{M}$ and $F_{21cm}$.

important to note the large uncertainty associated with these predicted mass-loss rates. This is due to the large uncertainty in the stellar masses obtained through the calibration of Martins, Schaerer, and Hillier (2005) and Searle et al. (2008) for the O and B stars respectively. The uncertainty in the stellar mass is ±50% and the resulting error in the calculated $\dot{M}$ values is represented by the shaded regions in Figure 4.4.

In order to predict the 21cm thermal free-free (smooth wind) flux as a function of spectral type for OI, III, V and BI stars, the derived $\dot{M}$ values as a function of spectral type could be used in conjunction with Equation 4.8. The equation requires the wind terminal velocity (as previously obtained from Prinja, Barlow, and Howarth 1990) and also the distance to the object (for which the distance to the Cygnus OB2 association $D = 1.4$ kpc, see Section 1.4.3 for further details, is used). For each star the Gaunt factor has been approximated to the free-free gaunt factor:

$$g_{ff} \approx 9.77 \left(1 + 0.13 \log \left(T_e^{3/2}/\nu\sqrt{\overline{(Z^2)}}\right)\right) \tag{4.12}$$

as defined by Leitherer and Robert (1991). It has been calculated assuming a constant relation between the stellar effective temperature and the electron temperature in the wind, $T_e = 0.5T_{eff}$ (Drew, 1989). Whilst the bound-free gaunt factor is also known to play a role, its contribution (at a wavelength of $\lambda = 21cm$) is found to be negligible and has thus been omitted from this calculation of the 21cm flux (see Table IV of Waters and Lamers 1984). For the purpose of the exercise demonstrated in Figure 4.4, the remaining parameters $\mu, \gamma$ and $\overline{Z^2}$ were taken to be 1.27, 1.0 and 1.0 respectively in the application of Equation 4.8. This assumes that Helium is singly ionised in the radio emitting regions of the winds for all stars with an abundance of $n_{He}/n_H = 0.1$.

The bottom panel of Figure 4.4, detailing the predicted 21cm flux as a function of effective temperature gives an insight into sensitivities required to detect the 21cm emission from the winds of early type stars in the Cyg OB association. As predicted by the mass-loss recipe Vink, Koter, and Lamers (2001), sensitivities between $\sim$ 1 - 20 $\mu$Jy are required to detect the 21cm radio emission for the majority of early type stars (with the exception of late OV stars). It is important to note however, that these predicted fluxes assume the

emission is entirely thermal in origin and furthermore that $f_{cl} = 1$, i.e. the wind is entirely smooth and free from small-scale structure. Any non-thermal emission would increase the 21 cm flux received by an observer, similarly the presence of wind structure in the radio emitting region would also increase the detected flux since $F_\nu \propto \dot{M}^{4/3} \propto f_{cl}^{2/3}$.

## 4.3 Mass-loss rates from single massive stars in Cyg OB2

Exploiting the imaging of single massive stars within the Cyg OB2 association from the COBRaS L-band observations, the 21cm flux densities of B supergiant and O3 to O6 supergiant and giant stars are constrained. These fluxes are then translated into 'smooth-wind' mass-loss upper limits for each of the stars within the sample. Furthermore, the first ever resolved 21cm detections of the hypergiant (and luminous blue variable (LBV) candidate) Cyg OB2 #12 are discussed for multiple observations (from Apr. 11th and Apr. 26th of 2014), to look for any variation in its flux density over the 15 day period.

### 4.3.1 Single massive star sample selection

The mass-loss rates of single massive stars within the Cygnus OB2 association can be derived using Equation 4.8 (re-arranging and solving for $\dot{M}$) and the flux measurements from the COBRaS L-band dataset. The primary set of observations were acquired on the 2014 Apr. 26th (see Table 2.3 for a full log of the COBRaS L-band observations) and their subsequent reduction and calibration is described throughout Section 2. The calibrated data produced images with a sensitivity between $\sim$ 20 - 25 $\mu$Jy within the centre of each of the seven target field pointings. With reference to the bottom panel of Figure 4.4 (i.e. the predicted 21cm fluxes from theoretical $\dot{M}$ values) we choose to limit our sample to those stars expected to have the densest winds, namely any O I, O III or B I stars within our field of view.

Starting from the most recent census of massive stars within the Cyg OB2 association from Wright, Drew, and Mohr-Smith (2015), we chose only stars within the field of view covered by the COBRaS L-band observations and of a O I-O III or B I spectral classification. The

sample was further limited to those stars known to be single, or if they are in a binary system, have sufficient separation from their companion that any wind-wind interaction can be considered as negligible, i.e. the expected emission is entirely thermal in origin. This assumption is crucial in deriving both the radio mass-loss rates and investigating wind structure (see e.g. Blomme et al. 2003). The presence of one or more companion stars will facilitate the production of non-thermal (synchrotron) emission within the colliding wind region(s) (e.g. as has been shown with the O6I+O5.5III(f) binary Cyg OB2 #8A; Blomme et al. 2010). As previously mentioned, the presence of non-thermal emission will contribute to the 21cm flux received, leading to an over-estimate of the derived mass-loss rate. For further details on the non-thermal emission of massive star binaries, please see Chapter 5.

Table 4.1 includes the final target sample of the nine single massive stars used for this study. Each star was imaged individually from the COBRaS L-band observations taken on Apr. 26th using the AIPS task IMAGR. Each 512 × 512 image was produced using a cell (pixel) size of 0.04 arcseconds, under a natural weighting scheme, and went through 500 iterations of the CLEAN algorithm. The sample includes two early O supergiant stars, two mid O giant stars, and five early B supergiant stars, one of which is a candidate LBV.

### 4.3.2 Adopted fundamental parameters

Of the nine stars within the target sample, the candidate LBV star Cyg OB2 #12 was the only detection within the COBRaS L-band observations. For the remaining sources a strict $3\sigma$ ($3\times$ the image RMS) upper limit was placed upon their 21cm flux densities. Each limit was applied to Equation 4.8 (taken from Wright and Barlow 1975) in order to obtain smooth-wind mass-loss upper limits for these objects (as is shown in Table 4.1).

In each case, the distance $D$ used in the application of Equation 4.8 was taken to 1.4 kpc (for further information see Section 1.4.3). The remaining parameters however, were selected according to the fundamental parameters of each star. Where possible, the stellar and wind parameters ($T_{eff}, v_{\infty}, \log g, M_{\star}$) were chosen from a full nLTE analysis. For both Cyg OB2 #7 and #8C, these were taken from Mokiem et al. (2005) and were derived from

TABLE 4.1: COBRaS L-band measured flux densities for the sample of stars. For Cyg #7, 8C and 12 the $T_{eff}$, $v_\infty$, log$g$ and $M_{spec}$ values are taken from the literature where a full nLTE analysis has been conducted in their derivation. Typical errors on these stellar parameters are $\Delta T_{eff} = \pm 500 - 1000$K. $\Delta v_\infty = \pm 50 - 100 \mathrm{kms}^{-1}$, $\Delta \log g = \pm 0.1 - 0.38$dex and $M_{spec}$ is uncertain between 35% and 50%. For the remaining stars the parameters are adopted from standard spectral type values taken from the references listed, note that parameters taken from references 4 and 5 were derived from calibrations. Predicted $\dot{M}$ values are calculated using the prescription from Vink, Koter, and Lamers (2001) with a revised metallicity value, Z = 0.013 following Asplund et al. (2009). We denote our derived mass-loss rate as $\dot{M}_{max}$ since we have set $f_{cl} = 1$.

| RA (J2000) | DEC (J2000) | S58 | MT91 | Other | Spectral Type | $T_{eff}$ (K) | $v_\infty$ km s$^{-1}$ | log$g$ | $M_{spec}$ $M_\odot$ | Flux Density ($\mu$Jy) | $\dot{M}_{max}$ $10^{-6}\,M_\odot \mathrm{yr}^{-1}$ | Predicted $\dot{M}$ $10^{-6}\,M_\odot \mathrm{yr}^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 32 40.88 | 41 14 29.3 | 12 | 304 | - | B3.5Ia+ | 13700 [1] | 400 [1] | 1.70 [1] | 110 [1] | 1013±55 | 5.4±1.4 | 24.5 |
| 20 33 14.16 | 41 20 21.5 | 7 | 457 | - | O3If | 45800 [2] | 3080 [3] | 3.94 [2] | 65 [2] | <72 | <4.8 | 3.5 |
| 20 33 18.02 | 41 18 31.0 | 8C | 483 | - | O5III | 41800 [2] | 2650 [3] | 3.74 [2] | 49 [2] | <71 | <4.1 | 1.9 |
| 20 33 08.78 | 41 13 18.1 | 22 | 417 | - | O3If | 42551 [4] | 3150 [6] | 3.73 [4] | 67 [4] | <61 | <4.4 | 4.3 |
| 20 33 14.84 | 41 18 41.4 | 8B | 462 | - | O6.5III | 35644 [4] | 2545 [6] | 3.63 [4] | 34 [4] | <78 | <4.3 | 0.7 |
| 20 32 39.06 | 41 00 07.8 | - | - | E47 | B0Ia | 28100 [5] | 1535 [6] | 2.99 [5] | 25 [5] | <87 | <2.9 | 0.8 |
| 20 33 39.14 | 41 19 26.1 | 19 | 601 | - | B0Iab | 28900 [5] | 1535 [6] | 3.13 [5] | 31 [5] | <63 | <2.2 | 1.1 |
| 20 33 30.81 | 41 15 22.7 | 18 | 556 | - | B1Ib | 21700 [5] | 1065 [6] | 2.67 [5] | 22 [5] | <73 | <1.8 | 1.6 |
| 20 33 33.97 | 41 19 38.4 | - | 573 | - | B3I | 16400 [5] | 590 [6] | 2.16 [5] | 19 [5] | <58 | <0.8 | 1.4 |

References: **1** Clark et al., 2012, **2** Mokiem et al., 2005, **3** Herrero et al., 2001, **4** Martins, Schaerer, and Hillier, 2005, **5** Searle et al., 2008, **6** Prinja, Barlow, and Howarth, 1990.

using the radiative transfer code FASTWIND (Santolaya-Rey, Puls, and Herrero, 1997). The adopted wind terminal velocity, $v_\infty$ values for these two O-type stars (as listed in Mokiem et al. 2005), were first derived by Herrero et al. (2001) in their study of HST STIS UV spectra. For Cyg OB2 #12, the stellar and wind parameters were obtained from an analysis of UV-radio photometric and spectroscopic data using CMFGEN (Hillier and Miller, 1998) conducted by Clark et al. (2012). For the remaining six stars however, the stellar and wind parameters were obtained through the calibrations of Martins, Schaerer, and Hillier (2005) and Searle et al. (2008) for the O and B stars in the sample respectively (as described in Section 4.2.4).

The remaining parameters ($\mu, \gamma, \overline{Z^2}, g_\nu(T)$) describe the ionisation conditions of the wind. To a good approximation (for small $\nu$), the Gaunt factor reduces to the free-free Gaunt factor, $g_{ff}$ (Equation 4.12) as described by Leitherer and Robert (1991). This is calculated through using the electron temperature in the wind, which relates to the stellar effective temperature as $T_e = 0.5T_{eff}$ (Drew, 1989). The mean atomic weight of gas, $\mu$, the ratio of the electron to ion density, $\gamma$, and the mean ionic charge, $\overline{Z^2}$, all depend on the ionisation state of the wind. Hydrogen is expected to be fully ionised within these hot star winds, yet the ionisation state of helium depends upon $T_{eff}$, the radial distance and the wind density (Morford et al., 2016). The nLTE radiative transfer code CMFGEN, used to fit the multi-wavelength (UV to radio) observations of Cyg OB2 #7, suggests that $He^+$ dominates over $He^{2+}$, even in the most favourable scenario of vanishing clumping in the outer wind (Najarro et al. 2008; Najarro, Hanson, and Puls 2011 and F. Najarro, private communication). Note, if any wind clumping were to exist within the radio region, the over density would seek to enhance recombination favouring the formation of $He^+$ over $He^{2+}$. Helium is therefore adopted to be singly ionised for all stars in the sample with the exception of Cyg OB2 #12 ($T_{eff}$ = 13.8 kK), where neutral helium is expected to dominate in the radio region (Clark et al. 2012, F. Najarro, private communication, see also Morford et al. 2016 for further details). Thus assuming a helium abundance of $n_{He}/n_H = 0.1$ across all sample stars, we adopt values of $\mu = 1.27$ (1.4 in the case of Cyg OB2 #12), $\gamma = 1$ and $\overline{Z^2} = 1.0$. Notes on individual stars are provided in the following sections.

### 4.3.3 Cygnus OB2 #7

This O3If star has previously been extensively observed and modelled with nLTE codes such as FASTWIND and CMFGEN (see Herrero, Puls, and Villamariz, 2000; Herrero, Puls, and Najarro, 2002; Herrero, Najarro, and Puls, 2003; Mokiem et al., 2005; Puls et al., 2006; Najarro, Hanson, and Puls, 2011; Maryeva and Zhuchkov, 2012; Maryeva, Klochkova, and Chentsov, 2013). Despite being part of previous radio surveys (Bieging, Abbott, and Churchwell, 1989; Setia Gunawan et al., 2003) it is yet to be detected at radio wavelengths, with a previous $3\,\sigma$ flux limit of 1.0 mJy at L-band (21cm) (Setia Gunawan et al., 2003).

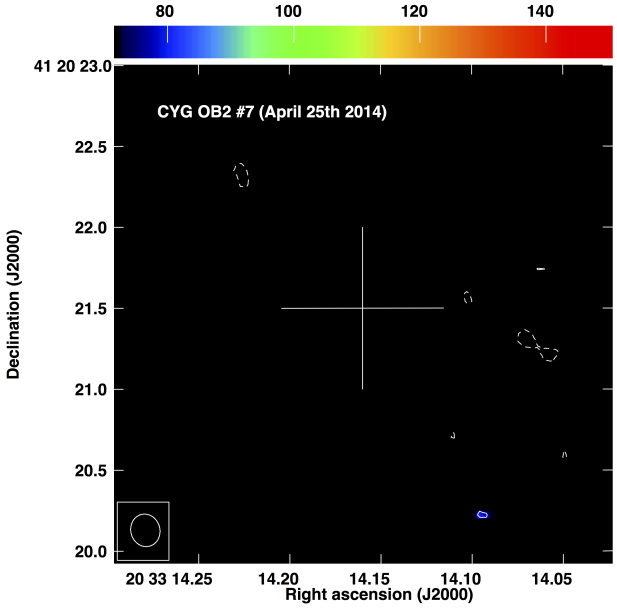



FIGURE 4.5: A blank field image of Cyg OB2 #7, image $1\sigma$ rms = 24 $\mu$Jy/beam. Upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms. Image taken from Morford et al. (2016).

Despite vastly improved sensitivities with e-MERLIN, the COBRaS L-band observations provide only an upper limit for Cyg OB2 #7. However, the observations reported here improve upon previous L-band (21cm) upper limit flux densities by over a factor of $\sim$10, reaching a $3\,\sigma$ flux density of 72 $\mu$Jy. This corresponds to an upper limit on the mass-loss rate of Cyg OB2 #7 of $4.8{\times}10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$. In comparison to the most recent values obtained via nLTE analysis of the H$\alpha$ line, our result is approximately a factor of two lower (e.g. Maryeva, Klochkova, and Chentsov, 2013). Using standard stellar parameters as a

function of spectral type (see Martins, Schaerer, and Hillier, 2005) and the recommended mass-loss recipe from Vink, Koter, and Lamers (2001) to calculate a predicted $\dot{M}$ for Cyg OB2 #7, provides values in good agreement with our result (see Table 4.1).

The first full nLTE analysis was carried out on Cyg OB2 #7 (Herrero, Puls, and Villamariz, 2000), using the model atmosphere code FASTWIND. H$\alpha$ and HeII$\lambda$4686 in combination with other optical lines were used to sample the inner-most wind regions ($R \lesssim 3\ R_\star$; see e.g. Prinja, Fullerton, and Crowther 1996) to derive a value of $\dot{M} = 11.2{\times}10^{-6}\,M_\odot yr^{-1}$ ($8.3{\times}10^{-6}\,M_\odot yr^{-1}$ at 1.4kpc). Furthermore, the same authors used an updated version of the FASTWIND code to account for both metal-line blocking and blanketing, that led to the derivation of similar $\dot{M}$ values (Herrero, Puls, and Najarro, 2002; Herrero, Najarro, and Puls, 2003). Mokiem et al. (2005) combined the FASTWIND code with a genetic algorithm based optimisation routine known as PIKAIA to automate the spectrum fitting process and again derived a 'smooth wind' mass-loss rate of $\sim 10{\times}10^{-6}\,M_\odot yr^{-1}$ ($7.4{\times}10^{-6}\,M_\odot yr^{-1}$ at 1.4kpc).

Subsequent studies of Cyg OB2 #7 were carried out by Puls et al. (2006), who conducted a multi-wavelength analysis incorporating radio, infrared (IR) and H$\alpha$ observations. By probing these different wind regions, they were able to put constraints on the radial stratification of the clumping factor. Their multi-wavelength approach, constrained by non-detections in VLA observations at 6 and 3.5 cm, derived an upper limit on $\dot{M}$ comparable to that found here and in Morford et al. (2016). For completeness, please note that Puls et al. (2006) assumed a Helium enrichment $Y_{He} = 0.21$, contrary to the value used here (of $Y_{He} = 0.1$) and in Morford et al. (2016). This however has a negligible effect on the assumed ionisation conditions and therefore the resulting mass-loss rate. Assuming the free-free thermal flux scales with frequency as $\nu^{0.6}$, the 21cm upper limit found here can be translated into a flux of 145 and 202 $\mu$Jy at 6 and 3.5 cm respectively, i.e. consistent with those derived by Puls et al. (2006). The same authors fixed the outer-wind clumping factor $f_{cl} = 1$ to derive a maximum mass-loss for Cyg OB2 #7. Whilst the results found here support this notion and provide good evidence that the outer wind regions are less clumped than the inner (H$\alpha$) wind regions, we stress that the clumping factor cannot be fully constrained at the radio photosphere without explicit knowledge of the star's true

mass-loss rate.

Further nLTE analysis utilising the H$\alpha$ line diagnostic (see Najarro, Hanson, and Puls, 2011; Maryeva and Zhuchkov, 2012; Maryeva, Klochkova, and Chentsov, 2013), all derive a 'smooth wind' mass-loss rate consistent with previous nLTE analysis of the star (i.e. are around a factor of two larger than found here). In the case of Cyg OB2 #7, the discrepancy between inner-wind region (H$\alpha$) and outer-wind region (radio) mass-loss rates is clear and must be attributed to the effect of wind structure (Morford et al., 2016).

### 4.3.4 Cygnus OB2 #8C

Previously considered an Oif type star (Massey and Thompson, 1991), Cyg OB2 #8C has since been re-classified as an Oiii by Kiminki et al. (2007). As in the case of Cyg OB2 #7, this star is yet to be detected at radio wavelengths. Whilst these COBRaS data provide the most sensitive radio observations of the Cygnus OB2 association to date, they provide only a 3$\sigma$ upper limit of 71 $\mu$Jy to the 21cm flux density of Cyg OB2 #8C. This corresponds to an upper limit on the mass-loss rate of 4.1 $\times 10^{-6}$ $M_{\odot}yr^{-1}$ (Morford et al., 2016).

There exists a multitude of nLTE spectroscopic analysis within the literature based upon Cyg OB2 #8C. Herrero, Puls, and Najarro (2002), utilising FASTWIND to model its stellar spectrum, calculate its mass-loss rate to be $2.3\times10^{-6}\,M_{\odot}yr^{-1}$ ($1.7\times10^{-6}\,M_{\odot}yr^{-1}$ at 1.4kpc) and later find a lower value of $1.7\times10^{-6}\,M_{\odot}yr^{-1}$ ($1.3\times10^{-6}\,M_{\odot}yr^{-1}$ at 1.4kpc; Herrero, Najarro, and Puls, 2003). Similarly, Mokiem et al. (2005) used FASTWIND to calculate a mass-loss rate of $3.4\times10^{-6}\,M_{\odot}yr^{-1}$ ($2.5\times10^{-6}\,M_{\odot}yr^{-1}$ at 1.4kpc). The multi-wavelength approach taken by Puls et al. (2006) utilised a 200 $\mu$Jy radio upper limit taken from Bieging, Abbott, and Churchwell (1989) in their calculation. Although their analysis of Cyg OB2 #8C failed to fully constrain the mass-loss and clumping properties across all regions they quote a $\dot{M}$ upper limit of $4.3\times10^{-6}\,M_{\odot}yr^{-1}$ ($3.2\times10^{-6}\,M_{\odot}yr^{-1}$ at 1.4kpc) based primarily on the radio flux density upper limit. Furthermore they calculate a lower $\dot{M}$ value of $3.5\times10^{-6}\,M_{\odot}yr^{-1}$ ($2.6\times10^{-6}\,M_{\odot}yr^{-1}$ at 1.4kpc) based upon the model fit to the H$\alpha$ line. Each of these previously derived values are in broad agreement with the upper limit of the mass-loss rate derived here. It is clear that higher sensitivities at radio wavelengths are

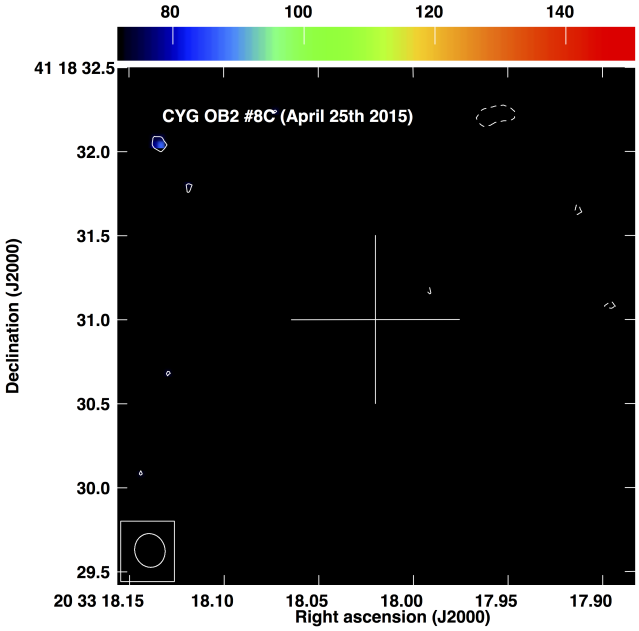



FIGURE 4.6: A blank field image of Cyg OB2 #8C, image $1\sigma$ rms = 24 $\mu$Jy/beam. Upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms. Image taken from Morford et al. (2016).

required to detect and constrain the mass-loss rate of early O-type giant stars at a distance of 1.4 kpc from the Earth.

### 4.3.5 Cygnus OB2 #12

There exists a rich 'zoo' of stellar objects between those on the main sequence (MS) and those of hydrogen depleted Wolf-Rayet (WR) stars. Such objects are blue hypergiants (BHGs), yellow hypergiants (YHGs), LBVs and P-Cygni supergiants. Unfortunately we lack a comprehensive theoretical framework for such stars. The uncertainty lies within the current difficulty to acquire an evolutionary scheme as a function of stellar mass that incorporates these exotic objects. Moreover, the current uncertainties in the mass-loss rate of early type stars, which dictates their evolutionary sequence, adds further weight to the issue. Hence observations of these objects are vital in order to investigate their properties and to test stellar evolutionary theory.

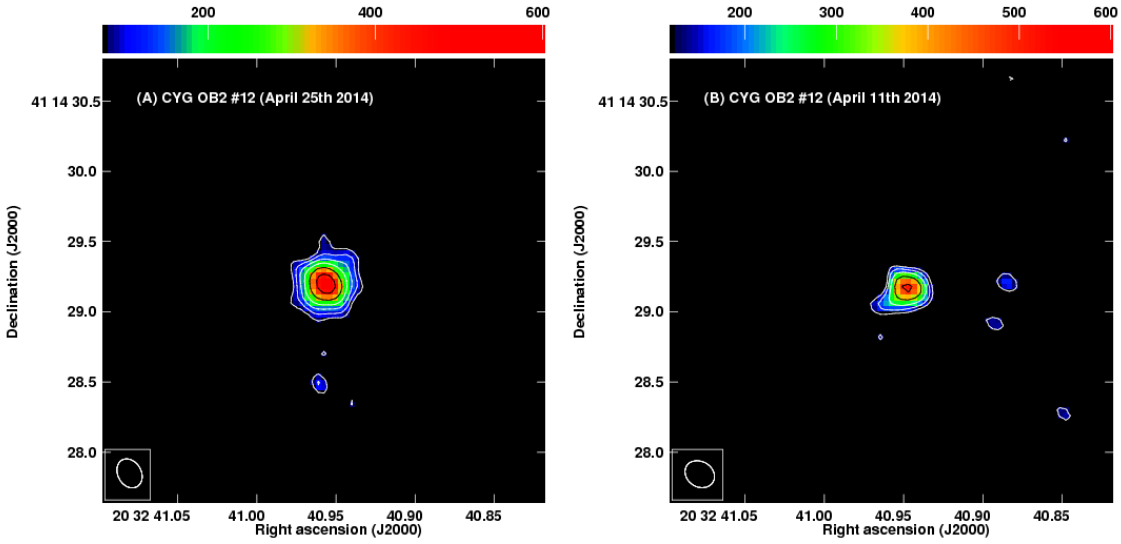



FIGURE 4.7: (A) Cyg OB2 #12 the first ever resolved image at 21cm, from the COBRaS observations taken on Apr. 26th, $1\sigma$ rms = 24 $\mu$Jy/beam; (B) a second image of Cyg OB2 #12 from COBRaS observations taken on Apr. 11th, $1\sigma$ rms = 40 $\mu$Jy/beam. The upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms. Images taken from Morford et al. (2016).

Cyg OB2 #12 (ID# 12 in Table 3.3) is a luminosity class Ia+ star (Keenan, 1971). Van Genderen (2001) later described this class as BHGs which not only differ from blue supergiants (BSG) by their large luminosity, but also spectroscopically with the presence of P Cygni balmer line emission. Cyg OB2 #12 has always been a particularly interesting BHG case due to its extremely high luminosity (Schulte, 1958). Whilst showcasing some properties of a LBV star, it is also missing some of the typical characteristics leading to uncertainties upon its exact classification (Clark, Larionov, and Arkharov, 2005; Clark et al., 2012).

LBVs are massive, unstable stars found in the upper left hand region of the HR diagram. Whilst having extremely high luminosities ($\sim 10^6$ $L_\odot$) and large mass-loss rates (up to $10^{-4}$ $M_\odot yr^{-1}$), they are found to be significantly variable both photometrically and spectroscopically (Humphreys and Davidson, 1994). They have been observed to have two types of variability, the first of which is reflected in their visual magnitude and is a result of how they cool and expand (or heat and contract), shifting to redder (bluer) colours. The

second is a consequence of significant mass-loss episodes such as the case for $\eta$ Carinae. These eruptions are far more rare, with only two known examples in our Galaxy (Clark, Larionov, and Arkharov, 2005).

Figure 4.7 shows the first ever resolved detections at 21cm of Cyg OB2 #12. Imaged from the COBRaS L-band observations taken on 2014 Apr. 26th, its measured flux density, $F_{21cm} = 1013 \pm 55$ $\mu$Jy (see Figure. 4.7A). Assuming a smooth wind model ($f_{cl} = 1$) and that the flux received is purely thermal free-free emission, we calculate its mass-loss rate to be $\dot{M} = 5.4 \pm 1.4 \times 10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$ from Equation 4.8. Clark et al. (2012) modelled Cyg OB2 #12 using CMFGEN to infer a value of $\dot{M} = 3.0 \times 10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$ ($2.1 \times 10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$ at 1.4 kpc), with a clumping factor $f_{cl} = 25$. Their value of $f_{cl}$ was derived using a modified version of CMFGEN's clumping prescription to account for the low terminal velocity of Cyg OB2 #12. Predominantly constrained using the H$\alpha - \beta$ and Br$\alpha$ emission components, the IR, sub-mm and radio continuum (Clark et al., 2012), this value of $f_{cl} = 25$ holds from R $>$ 40 R$_{\star}$ (F. Najarro, priv. comm.), covering the entire radio emitting region. The derived $\dot{M}$ translates into an 'unclumped' (smooth-wind) value of $15 \times 10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$ ($10.7 \times 10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$ at 1.4kpc), giving a discrepancy of a factor of $\sim 2$ in comparison to the value derived from the 21cm emission (Morford et al., 2016). As is the case with Cyg OB2 #7, this highlights the uncertainty in current $\dot{M}$ diagnostics and adds to the growing evidence for the disparity between H$\alpha$ and radio inferred mass-loss rates.

The uncertainty of this object and the ongoing debate upon its precise classification cannot be overlooked. The significantly lower than previously found $\dot{M}$ value derived here could instead be explained by the variability of the object. COBRaS L-band observations were also obtained on Apr. 11th, $\sim$15 days prior to the primary set of observations. With a sensitivity of $\sim 40$ $\mu$Jy, the observations taken on Apr.11th yielded a flux of $598 \pm 61$ $\mu$Jy (see Figure 4.7B) corresponding to a 'smooth-wind' mass-loss rate of $3.6 \times 10^{-6}\,\mathrm{M_{\odot}yr^{-1}}$ (Morford et al., 2016). This implies a 50% increase in the mass-loss rate of Cyg OB2 #12 (or a 69% increase in the flux density) over the 15 day period (Morford et al., 2016). A possible constraint on the origin of this variation may be derived from considering the 'effective radius' of the radio emission, defined as the radial distance at which the free-free optical depth is 0.244 (Wright and Barlow, 1975).

Using the stellar parameters found in Table 4.1 (and $R_\star = 246\ R_\odot$; Clark et al. 2012), the effective radius of 21cm emission in Cyg OB2 #12 is $\sim 86\ R_\star$ (Morford et al., 2016). To cover this distance at a constant velocity of $v_\infty = 400\ kms^{-1}$ would require $\sim 424$ days, i.e. much longer than the 15 days between the two observation epochs (Morford et al., 2016). For comparison a typical O-type star with $v_\infty = 2600\ kms^{-1}$, $T_{eff} = 40$ kK, $R_\star = 10\ R_\odot$ and a mass-loss rate of $4 \times 10^{-6}\ M_\odot yr^{-1}$ would need $\sim 9$ days to cover the distance of its 21cm effective radius ($\sim 304\ R_\star$; Morford et al. 2016). Thus, the observed variation in the 21cm flux of Cyg OB2 #12 cannot be explained by a global mass-flux variation.

Please note that the Apr. 26th observations are approximately 3× the on-source integration time than those of the Apr. 11th and the difference between the coverage in hour angle results in a different primary beam shape and size between the two observation epochs. The object is resolved in both epochs, with a deconvolved angular size of 285 milliarcseconds (mas) and 453 mas for the Apr. 11th and Apr. 26th observations respectively (Morford et al., 2016). Short term variations in the flux of Cyg OB2 #12 have also been found in previous radio observations. Bieging, Abbott, and Churchwell (1989) found a 70% variation at 6 cm in the flux of Cyg OB2 #12, whilst Scuderi et al. (1998) observed a 50% variation in its radio flux over the timescale of a month at 2, 3.6 and 6 cm. Furthermore, its X-ray flux has also been found to vary on the 10% level over timescales of up to a week (Rauw, 2011).

Cyg OB2 #12 has been extensively studied. Clark et al. (2012) report that the combination of its extremely high luminosity and low temperature imply its position on the HR diagram cannot be matched to any theoretical isochrone applied to its host association Cyg OB2. Cazorla, Nazé, and Rauw (2014) looked at XMM-Newton and Swift X-ray observations of Cyg #12 and find a marked decrease in X-ray flux in recent years (40% from 2004 to 2011), compatible either with a wind-wind collision in a wide binary or the aftermath of a recent eruption. No evidence for a companion star had previously been found until Caballero-Nieves et al. (2014) detected a close companion separated by an angular distance of 63.6 mas. Their finding has since been confirmed by Maryeva et al. (2016), who further resolved a very faint third counterpart. Note however that Caballero-Nieves et al. (2014) report that their detected secondary is too faint to substantially decrease the luminosity of Cyg

OB2 #12, and hence alter the conclusions drawn from Clark et al. (2012). Depending on the nature of Cyg OB2 #12's companions, this potentially undermines the assumption that the radio flux detected in the COBRaS observations presented here is completely thermal in origin. However, it is important to note that any non-thermal emission as a result of a colliding wind region would only further contribute to the 21cm flux received here, implying a smaller contribution from thermal free-free emission. Hence the presence of non-thermal emission would only seek to lower the mass-loss rate derived here, in contradiction to previous $\dot{M}$ estimates (e.g. Clark et al., 2012).

### 4.3.6 The remaining sample selection

The remaining six stars within the sample are Cyg OB2 #22, #8B, #19, #18, E47 and MT 573. For these objects, no previous nLTE modelling of their stellar spectra can be found in the literature, meaning comparisons to the derived mass-loss rate upper limits found in this study (and in that from Morford et al. 2016) must be made with predicted $\dot{M}$ values based upon calibrations. This prescription was found using the methods described in Section 4.2.4, whereby stellar calibrations as a function of spectral type have been fed into the mass-loss recipe from Vink, Koter, and Lamers (2001). The values are shown alongside those derived from the upper limit on the 21cm flux density in Table 4.1. With the exception of the O6.5III star, Cyg OB2 #8C and the B0Ia star E47, the $\dot{M}$ upper limits are within broad agreement (a factor $\lesssim 2$) to those predicted using the prescription from Vink, Koter, and Lamers (2001).

Many previous studies, across a multitude of wavebands (e.g. Najarro, Hanson, and Puls 2011: IR lines; Cohen et al. 2014: X-ray line emission; Sundqvist et al. 2011; Šurlan et al. 2013; Sundqvist, Puls, and Owocki 2014: UV-lines including velocity porosity and optical lines), consistently derive OB star mass-loss rates $\sim 2-3$ times smaller than those theoretically predicted from the Vink, Koter, and Lamers (2001) prescription. Thus, if this discrepancy between observed and theoretically predicted mass-loss rates is assumed, a clumping factor, $f_{cl} \sim 4$ would be necessary to pull the observed upper limit $\dot{M}$ values downwards and to re-create the previously found discrepancy.

Interestingly, the two B supergiant stars of MT573 and Cyg OB2 #12 have a derived $\dot{M}$ upper limit, lower than the theoretical prediction inferred from the Vink, Koter, and Lamers (2001) recipe by factors of 2 and 5 respectively. Coincidentally, these are the only two stars within the sample whose stellar effective temperatures lie beneath the observed bi-stability jump at 20 kK (see Evans et al. 2004; Crowther, Lennon, and Walborn 2006; Markova and Puls 2008; Petrov, Vink, and Gräfener 2016). Considering the rarity of radio observations of B supergiant stars, this result could provide crucial information regarding the on-going debate into the increase in $\dot{M}$ values as stars cross this bi-stability jump. For more information please see Vink, Koter, and Lamers (1999) and Markova and Puls (2008).

## 4.4 Discussion and conclusions

This chapter began with an introduction into the stellar winds of early type stars, including the observational diagnostics used to determine their wind properties and a deeper look in their driving mechanism by way of radiative line acceleration. In relaxing the assumption of a spherically-symmetric, homogeneous wind, the effects of small (clumping) and large (CIRs) scale wind structures were discussed in relation to the observed properties of such objects. It was shown how this has lead to the current uncertainties associated with mass-loss from massive stars and the profound effect such an uncertainty has within a greater astrophysical context. In order to resolve such an uncertainty, the 21cm e-MERLIN observations of Cyg OB2 from the COBRaS project have been used to investigate the stellar wind properties of a sample of nine massive stars.

The observations demonstrated that the flux densities for a sample of early O-type stars in the Cyg OB2 association are below $\sim 70\mu$Jy (see Table 4.1). Provided the emission is entirely thermal in origin and no (or minimal) small-scale structure (wind clumping) exists within the radio region, the upper limits on mass-loss rates of O3 I stars (i.e. the hottest and most luminous objects within the sample) were found to be between $\sim 4.4-4.8\times10^{-6}$ $M_{\odot}yr^{-1}$. Furthermore, the mass-loss rates of early B supergiant stars (B0 to B1) have been constrained to $\lesssim 1.8-2.9\times10^{-6}$ $M_{\odot}yr^{-1}$ (Morford et al., 2016). These $\dot{M}$ upper limits have

been shown to be broadly consistent[1] with those predicted from the theoretical predictions given by Vink, Koter, and Lamers (2001). Please note that the $\dot{M}$ predictions based on spectroscopic masses obtained through calibrations, as opposed to a nLTE analysis, are typically uncertain between 35% and 50% (see Figure 4.4).

If considering the 'effective radius' at 21cm of the early O-type stars within the sample, i.e. using the definition where the 21cm photosphere is at the point at which free-free optical depth $\tau_{ff} = 0.244$ (Wright and Barlow, 1975), then $R_{21cm} \sim 150R_{\star}$. Our calculations of the mass-loss rate from the 21cm emission at these regions in the stellar wind have assumed $f_{cl} = 1$, i.e. that no small-scale structure exists at this point in the wind. In reality however, this has been untested observationally, due to the high sensitivity required to detect the 21cm thermal free-free emission on $\mu$Jy scales. Therefore, from the upper limit on the 21cm flux density it is possible to infer i) a maximum 'smooth-wind' ($f_{cl} = 1$) mass-loss rate, or ii) a lower mass-loss rate for some value of $f_{cl} > 1$.

For stars in our sample such as Cyg OB2 #7, a comparison can be made to the nLTE spectroscopic analysis of Herrero, Puls, and Najarro (2002), Mokiem et al. (2005), Repolust et al. (2005), and Maryeva, Klochkova, and Chentsov (2013). Their line synthesis analysis of the H$\alpha$ and HeII lines, formed within the inner wind region ($\lesssim 3R_{\star}$), provide a 'smooth wind' mass-loss rate of $\sim 8.0 - 10 \times 10^{-6}$ $M_{\odot}yr^{-1}$. These are consistently high in comparison to the 21cm upper limit derived here, and provide strong evidence that the inner wind regions must be substantially more clumped ($f_{cl} \gtrsim 4$) in comparison to the outer, radio wind region sampled by the COBRaS L-band observations. This result is in agreement with the findings from Puls et al. (2006) whose multi-wavelength study of a sample of massive stars revealed a radial dependance on the clumping factor.

Porosity effects (macro-clumping) have been shown to reconcile UV resonance line inferred mass-loss rates, to those obtained through ($\rho^2$) recombination line diagnostics (Oskinova, Hamann, and Feldmeier, 2007; Sundqvist et al., 2011). However, porosity effects have only been found to act on optically thick line transitions. Furthermore, since the free-free opacity grows with $\lambda^2$, one should expect porosity effects to become important at long

[1]With the exception of the LBV candidate Cyg OB2 #12 and the two OIII stars within the sample.

wavelengths (Ignace, 2016). The work of Ignace (2016), who investigated the influence of macro-clumping on the spectral energy distributions (SEDs) of ionised winds, considered how the geometrical shape of the clumps would impact mass-loss rate determinations when taking into account porosity effects. Whilst the precise clump geometry is unknown, they are thought to expand in the azimuthal direction as they move away from the stellar surface, forming 'pancake-shaped' shell-like fragments. In this case, Ignace (2016) found that mass-loss determinations are unaffected by the inclusion of porosity effects.

An analysis of the 21cm observations presented here and in Morford et al. (2016) provide evidence of a radially stratified clumping factor in the winds of hot stars. Moreover, in comparison to nLTE H$\alpha$ line (inner wind) $\dot{M}$, diagnostics, these observations suggest that at least the inner wind regions are more heavily clumped than the outer, radio regions. The 1-dimensional hydrodynamical simulations of Runacres and Owocki (2002) show that the growth of the line-de-shadowing instability (LDI) predicts an increase in the clumping factor from the inner wind regions out to $\sim 50R_{\star}$. Depending on their precise details, their simulations predict clumps to remain in the stellar wind, far out in the radio region with values of $2.5 \lesssim f_{cl} \lesssim 6$ (Runacres and Owocki, 2002). Unfortunately, the single epoch radio continuum observations presented here do not provide any direct evidence of $f_{cl}$ at large radial distances in OB star winds and further observations are required, along with a robust estimate of the stars mass-loss rate, in order to fully constrain the radial dependance of $f_{cl}$.

### 4.4.1 A look ahead to the COBRaS C-band data

The COBRaS 6cm (C-band) observations are scheduled to commence from the beginning of 2017 onwards. With a total awarded allocation of 252 hours over 42 pointings and the recently upgraded e-MERLIN C-band receivers covering a full 2 GHz in bandwidth, the observations will provide $3\sigma$ flux densities down to a limit of $\sim 10$ $\mu$Jy. Covering the same area on the sky as those of the 21cm observations (see Figure 1.14), the 6cm observations will lead to the tightest constraints on the mass-loss rates derived from the thermal free-free radio emission in the outer winds of massive stars. Whilst the C-band observations

will offer an increased sensitivity in comparison to the 21cm observation presented here, the positive spectral index of thermal free-free emission, $F_\nu \propto \nu^\alpha$ with $\alpha \sim +0.6$, implies $F_{6cm} > F_{21cm}$ for these objects. This will lead to a greater number of detections of the single star population in the Cyg OB2 association, an increased sample size as a function of spectral type, accurate detected flux densities and tighter constraints on the stellar wind parameters of massive stars.

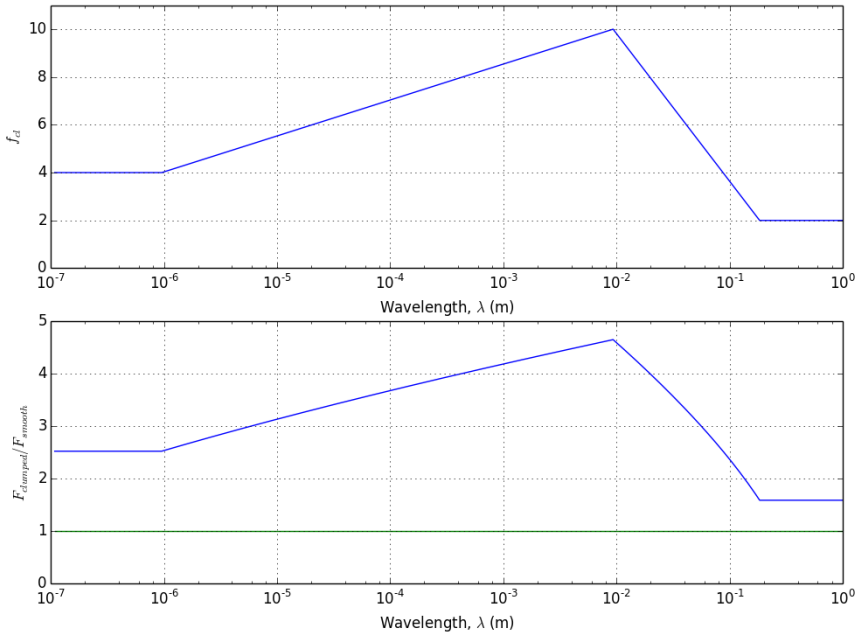



FIGURE 4.8: A semi-empirical model of the radial stratification of $f_{cl}$ as a function of the wavelength of free-free emission within the winds of hot stars. The bottom panel shows the ratio of calculated flux (using Equation 4.9) at the values of $f_{cl}$ shown in the top panel to that of a smooth wind flux ($f_{cl} = 1$). Values of $f_{cl}$ were chosen to be 4, 10 and 2 at wavelengths of 1$\mu$m, 1mm and 20cm respectively. Stellar parameters were chosen to represent that of a typical O-star, $T_{eff}$ = 40000K, $v_\infty$=2500, $\dot{M} = 2\times10^{-6}$, $\mu = 1.27$, Z=1 and $\gamma = 1$ (see Section 4.3.2 for further details). The green line in the bottom panel represents a smooth wind model ($fcl$=1).

Constructing a semi-empirical model of the radial stratification of the clumping factor as a function of wavelength (and hence in turn the radial distance in the wind) will enable a direct comparison to the future C-band detections of massive star winds in Cyg OB2. Figure 4.8 shows an example of such a model whereby three values of the clumping factor are chosen at three particular wavelengths, from which the flux density is calculated using Equation 4.9 as a function of $f_{cl}$ and $\lambda$ and compared to that of a smooth wind (i.e.

$f_{cl} = 1$). Whilst this model does not specifically relate the physical stratification of $f_{cl}$ as a function of the radial distance in the wind, it gives a good indication of such since the free-free radiation at increasing wavelengths is emitted at progressively larger radial distances in the wind. By tuning values of $f_{cl}$ at different wavelengths one is able to see the effect this has on the flux at a given wavelength, i.e. enabling a direct comparison to observations such as that shown in Figure 4.1 of the observations made of the B0Ia star $\epsilon$Ori. The COBRaS C-band dataset, the 21cm upper limits derived above, and observations at other wavelengths (e.g. in the mm and IR) are required to determine the amount of clumping as a function of radial distance (see e.g. Figure 4.1). Comparisons to semi-empirical models such as that shown in Figure 4.8 will then give a better understanding of small-scale structure in the winds of hots stars and ultimately lead to reliable mass-loss rates.

# Chapter 5

# 21cm radio emission from massive star binaries in Cygnus OB2

This chapter explores the incidence of non-thermal emission from the colliding winds of massive star binary systems in Cyg OB2. Utilising the 21cm (L-band) observations from the COBRaS Legacy project, the non-thermal radio emission associated with a number of detected (and previously well-studied) massive star binaries is presented. The flux densities of such systems are directly compared to previous observations as a function of the orbital phase, leading to the construction of their 21cm 'light curve'. Observations and analysis of this kind are crucial in developing reliable models of massive star binaries such that their orbital and stellar parameters can be constrained as a function of the radio flux associated with the non-thermal (synchrotron) radiation produced in their colliding wind region. Furthermore, it will be shown that observations at (long) wavelengths can be used to infer the possible presence of previously unidentified massive star binary systems. The radio observations of three potential new binary systems associated with the Cyg OB2 association are presented.

## 5.1 Massive star binaries

Massive stars are those considered to have a mass $\gtrsim 8$ $M_{\odot}$ and an effective temperature $\gtrsim 10^4$ K. These hot stars range in spectral type from O, B (and their respective evolved counterparts), to Wolf-Rayet (WR) stars. As discussed in Chapter 4, such massive stars

are known to produce a strong stellar wind which can carry away mass at a rate between $\sim 10^{-8}$ and $10^{-5}$ $M_{\odot}yr^{-1}$ and with a velocity between $\sim$ 1000 and 3000 $km\,s^{-1}$ . Whilst the works of Wright and Barlow (1975) and Panagia and Felli (1975) showed that such objects emit thermal free-free emission from the IR to radio wavelengths, observations throughout the 1980s began to show that these objects were also capable of producing non-thermal (synchrotron) emission (e.g. White 1985; see Section 5.2.2 for further details). The presence of synchrotron radiation implied that these stars were capable of accelerating particles up to relativistic speeds (White and Becker, 1983; Abbott, Bieging, and Churchwell, 1984; White, 1985). This gave rise to the realisation that massive stars were now sources of extremely energetic (relativistic) particles and potentially large contributors to the observed number of cosmic-rays within our Galaxy (De Becker and Raucq, 2013).

The production of non-thermal emission observed in early type stars requires the presence of relativistic electrons in an electromagnetic field. Several mechanisms, whereby shocks are able to produce particles at such high speeds have been proposed. Firstly, in the single star case, hydrodynamical shocks have been shown to be a consequence of line de-shadowing instabilities (LDI) intrinsic to the stellar winds of early type stars (Owocki and Rybicki, 1985; Owocki, Castor, and Rybicki, 1988; Feldmeier, 1995). These instabilities, formed within the radiatively driven winds, have also been shown to create wind structures ('clumping', see Section 1.2.2). It was proposed that these hydrodynamical shocks could accelerate electrons to relativistic speeds, and in the presence of a local magnetic field, lead to the production of synchrotron emission (De Becker, 2007). A second case however was proposed, by considering that several of these early type stars observed to exhibit non-thermal radiation, were actually part of a massive star binary system, i.e that shocks (and hence relativistic electrons) were produced from the collision between the stellar winds of the two binary members (see Section 5.2.2 for further details). Each stellar wind interacts with the other in a so called wind collision region (WCR) at a point, somewhere between the two stars, that is confined between two hydrodynamic shock fronts (Stevens, Blondin, and Pollock, 1992; Pittard and Stevens, 1997). For the remainder of this chapter, the work presented is limited to the non-thermal emission produced from the acceleration of particles within a magnetic field, from the colliding winds between two (or more) massive

stars (i.e. as described by De Becker and Raucq 2013 as particle accelerating colliding wind binaries, from here on PACWB). As discussed in Pittard and Dougherty (2006), the scenario of synchrotron radiation emitted from winds of single massive stars (due to LDI) is unlikely, due to the significantly large optical depths found within their winds which readily absorb the non-thermal emission. Only at sufficient radial distances in the wind (more than a few hundred radii; Pittard and Dougherty 2006) will the optical depth drop such that synchrotron emission could be readily observed from these objects.

In the collision region between the winds of two or more massive stars, the temperature of the gas (or plasma) is heated to a few $10^7$ K and leads to the production of thermal X-ray emission (De Becker, 2007). Comparatively, the plasma found within the ionised wind of either binary component will generally be on the order of $\sim 10^6$ K (at most). Therefore the WCR between two massive stars will be brighter in X-rays in comparison to the individual stars (De Becker, 2007). If assuming an empirical relationship between the X-ray emission and bolometric luminosity as a function of spectral type, any X-ray excess (along with the presence of non-thermal emission at radio wavelengths) will also be indicative of a WCR between two massive binary stars. Such systems (PACWBs) have also been predicted to emit non-thermal high energy photons (X-rays and $\gamma$-rays) through inverse compton scattering (Pittard and Dougherty, 2006; Palacio et al., 2016). The detection of $\gamma$-ray emission is limited to only the objects $\eta$ Carina (Tavani et al., 2009; Farnier, Walter, and Leyder, 2011) and WR 11 (Pshirkov, Vasiliev, and Postnov, 2016) and similarly hard X-rays have only been detected in a select few cases. Interestingly, the presence of such high energy photons has been found in HD159176, a very short period binary (3.4 days) that is not known to be a non-thermal radio emitter (Dougherty and Williams, 2000). This has been explained by the fact that close period massive star binaries are thought to experience significant free-free absorption within their stellar winds (Dougherty and Williams, 2000), hence the non detection of non-thermal radio emission in HD159176.

### 5.1.1 The impact of massive star binaries

The study of non-thermal radio emission from the WCR of massive star binaries is limited to a select number of examples. Three of such objects however, namely Cyg OB2 #5, Cyg OB2 #8A and Cyg OB2 #9, reside in the Cyg OB2 association and within the field of view covered by the COBRaS L-band Legacy observations. Despite this limited number, the majority of massive stars are estimated to exist within a binary (or multiple) system. Within the field population of massive stars in the solar neighbourhood, Mason et al. (1998) estimated that $\sim 70\%$ are found to be a part of a binary system. Sana and Evans (2011) and Sana, James, and Gosset (2011) and Sana et al. (2012) found a binary fraction of $\sim$50% for massive stars within young open clusters. Furthermore, for the Cyg OB2 association, Kobulnicky et al. (2014) found a binary fraction of 55% for those systems with an orbital period, $P < 5000$ days (please see Table 1.3). The high fraction of massive star binaries and their abundance throughout the observable universe cannot be neglected and highlights the necessity for their study as individual objects in order to understand their impact on the surrounding environment.

The impact of binary or multiple star systems on their host clusters has often been overlooked (Geller et al., 2015). From both an observational and modelling point of view this is perhaps understandable. Practically they are difficult to observe and including them in $N$-body simulations of stellar clusters comes at the cost of computational time. In reality however, the presence of binaries has been shown to alter a number of cluster properties and characteristics, as shown in observations and from the modelling of such systems. In estimating the dynamical mass of a cluster using the virial theorem for example, Kouwenhoven and De Grijs (2008) showed that this is often overestimated by up to an order of magnitude if ignoring the presence of binary systems which increases the measured velocity dispersion of the cluster. Within analytical evolutionary models of stellar clusters (Hills, 1980), it was shown that binaries act to increase the size of the cluster core if the density of the core is sufficiently low (such as that found in a typical open cluster). In this respect, binaries act as a source of energy, opposing the collapse of the cluster core and ultimately bearing influence on the evolution of the entire cluster system.

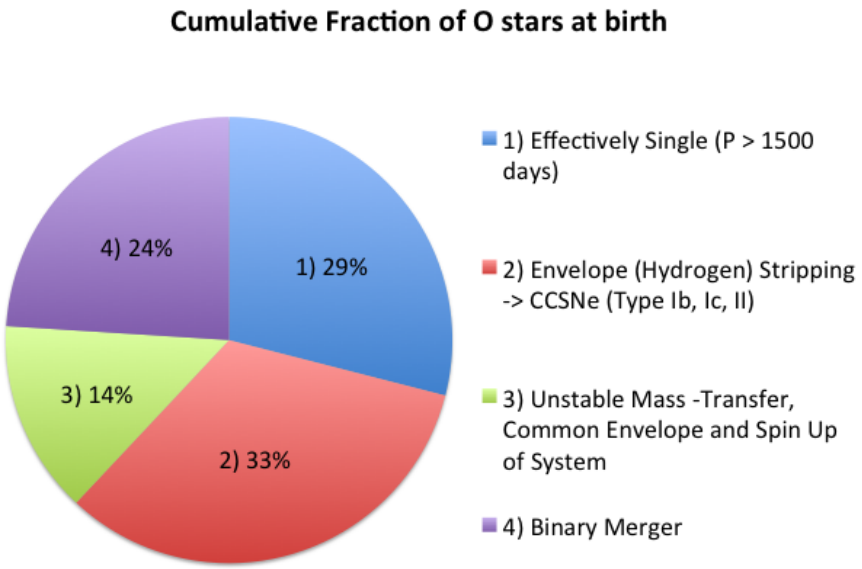



FIGURE 5.1: To illustrate the derived relative fractions of the various binary interaction scenarios of a sample of O stars as found by Sana et al. (2012).

From the point of view of the binary system itself, the evolutionary tracks of the two (or more) stars are also altered due to the interactions between the two stars. Whilst many binaries are quite wide systems where each component can be approximated as a single star (i.e. their interaction with one another does not affect their evolution) there still exists a large fraction of systems with a period, $P \lesssim 10$ yr which are close enough for mass to be transferred from one star to the other (Podsiadlowski, 2001). Analysing over 1800 spectra of over 71 single or multiple O-type objects, Sana et al. (2012) find a binary fraction of $0.69 \pm 0.09$ for their sample when taking a Monte-Carlo approach to simulate their observational biases. Based upon the detailed binary evolutionary calculations of Podsiadlowski, Joss, and Hsu (1992), Pols and Marinus (1994), Wellstein, Langer, and Braun (2001), and De Mink, Pols, and Hilditch (2007) they determine the relative frequencies of the proposed binary interaction scenarios for all the systems in their sample. For binaries with an orbital period less than 1500 days, Sana et al. (2012) find that mass-transfer plays a significant role in the subsequent evolution of either star and hence their ultimate fate. Figure 5.1 shows the relative fractions they find for the various interaction scenarios. $71 \pm 8\%$ of all O-type stars (counted individually from their sample) were found to be members of binary systems that will interact by Roche lobe overflow (RLOF; see Paczynski (1971)

and Eggleton (1983)). This is a mechanism whereby mass can be transferred between two stellar companions and depends upon their separation, respective radii and mass ratio $q = M_1/M_2$.

It is therefore the nature of the massive star binaries (i.e. their separation, orbital period and mass ratio) that can have a crucial effect on the evolution of the system as a whole, and indeed of each individual component. This in turn has a profound effect on stellar population models, severely altering the relative fraction of core collapse supernovae (CCSNe; of type Ib, Ic, II) and leads to the production of eruptive events due to binary mergers. Massive star binaries undoubtedly play a large role in shaping their environment and simply cannot be neglected in a wider astrophysical context. The present study into the non-thermal radio emission from massive star binaries aims to improve our understanding of such systems by unravelling the dependance of the binary characteristics on the radio flux density emitted from the WCR and obtaining a further handle on the fundamental binary fraction.

## 5.2 Non-thermal radio emission from massive star binaries

Particle accelerating colliding wind binary (PACWB) systems constitute two (or more) stars, each of an ample temperature and mass to harbour a stellar wind, and are of sufficient separation from one another that their winds collide. The collision of these energetic particles leads to the production of hydrodynamical shocks, in which particles (electron and ions) are accelerated to relativistic speeds. If such charged particles, travelling at relativistic speeds (with kinetic energy, $E_k >> m_e c^2$) are in the presence of a magnetic field, they will be accelerated along a spiral path following the line traced by the magnetic field. As the charged particles are constantly changing direction they are in effect accelerating and emitting radiation in the form of a photon with an energy dependent on the speed of the electron at the time of emission. The radiated photons are 'beamed' (confined) to a narrow cone in the direction of motion of the particle at the time of emission and polarised in the plane perpendicular to the magnetic field. For such synchrotron emission to be readily observed, the relativistic electrons must reside in an optically thin region

of the stellar wind(s), i.e. more than a few hundred stelar radii (Pittard and Dougherty, 2006). The implications of which span two-fold, firstly that the WCR must be located at a radial distance far away from either star such that free-free absorption does not halt the emission of synchrotron radiation and secondly (as aforementioned) this undermines the prior suggestion that non-thermal synchrotron radiation is a consequence of hydrodynamic shocks introduced by LDI (Van Loo, Runacres, and Blomme, 2005).

### 5.2.1 Synchrotron radiation

Synchrotron emission, discovered in 1947 by Elder et al. (1947) was first detected in an astronomical context within the jet associated with the elliptical galaxy, Messier 87 (Burbidge, 1956). The non-thermal emission has now been detected from many different astronomical sources including Active Galactic Nuclei (AGN), star forming galaxies, from the magnetosphere of Jupiter, quasars and the WCR of massive star binaries (see Sokolov and Ternov 1966). As previously mentioned, the two fundamental properties for its production are relativistic charged particles and a magnetic field. Depending on the relative strengths of these properties (i.e. magnetic field strength and the energy of the electron) synchrotron emission can occur at a huge range of wavelengths from the X-ray and ultraviolet, through to the visible and perhaps most commonly at radio wavelengths.

The spectrum of synchrotron emission follows a power-law decay and results from a summation of the emission spectra of the individual charged particles (usually electrons). This is highlighted in Figure 5.2 where the synchrotron spectrum of a single particle (shown in the top right hand corner) is created as it spirals around the magnetic field emitting photons at a range of frequencies peaking at the critical frequency, $\nu_0$. As the particle spirals around the magnetic field line, it loses energy and creates subsequently smaller spirals. With this energy loss, the critical frequency, $\nu_0$ of the photon also decreases, leading to the production of longer wavelength emission. Summing the contribution of the emission from each individual contribution (particle) leads to the characteristic synchrotron spectrum shown in Figure 5.2. The flux has a power-law like dependance on the observed frequency, $F_\nu \propto \nu^\alpha$ where the spectral index, $\alpha$ is typically found to be negative for synchrotron

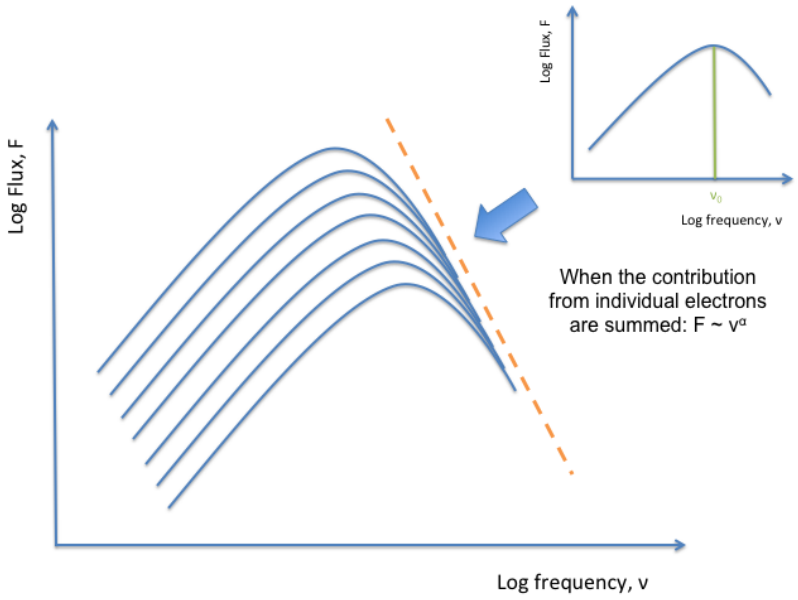



FIGURE 5.2: A schematic illustration of the characteristic synchrotron emission spectrum found by summing those of the individual contributions. Figure based on that from the online COSMOS - The SAO Encyclopaedia of Astronomy webpages created by the Swinbourne University.

emission, although its precise value is dependent on the physical properties leading to the emission.

The non-thermal emission produced from the WCR of a massive star binary is considered to be 'non-thermal' since it bears no dependance on the temperature of the emitting material. The spectral index, $\alpha$, of synchrotron emission produced from the collision between two massive star winds has generally been found to be between $\sim -0.5$ and $-1.0$ Bell (1978) and Pittard and Dougherty (2006). It has at times been stated as being flat, (i.e. $\alpha = 0$; Abbott, Beiging, and Churchwell, 1981; Abbott et al., 1986; Dougherty and Williams, 2000; Reimer, Pohl, and Reimer, 2006; Palacio et al., 2016), however the negative nature of the spectral index of the non-thermal emission is well below than the $\alpha \sim +0.6$ as predicted from the free-free thermal emission within the stellar winds of single stars (Wright and Barlow, 1975). The negative (or flat) spectral index associated with the WCR of massive star binaries implies that the COBRaS 21cm observations should be more sensitive to such emission, which increases towards longer wavelengths in comparison to the thermal free-free emission which drops off with increasing $\lambda$.

### 5.2.2 Colliding wind regions

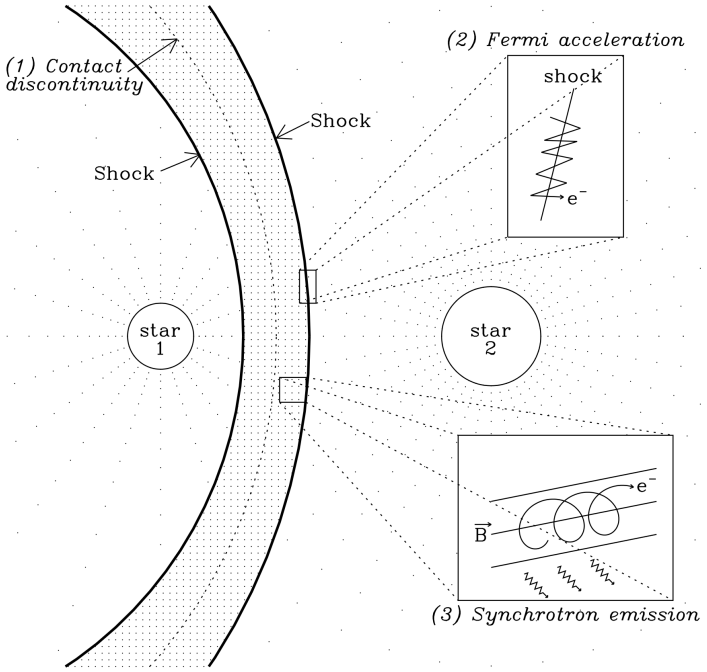



FIGURE 5.3: A schematic illustration of a wind collision region (WCR) formed between two massive stars within a binary system. Figure taken from Blomme (2011).

The presence of non-thermal synchrotron radiation can be explained by the presence of relativistic particles (electrons) moving in a local magnetic field. It has since been proposed that both of these elements can be found within PACWB binary systems. The first question is then how these charged particles (electrons, protons and even heavier nuclei) can be accelerated to speeds close to the speed of light. Figure 5.3 shows a schematic illustration of the WCR between two massive stars. It highlights how a fraction of the particles at either of the strong shock fronts are accelerated up to relativistic speeds through the first-order Fermi mechanism as described by Bell (1978). Through this mechanism, the charged particles (electrons) are bounced back and forth along the shock front (as shown in Figure 5.3), each time gaining energy from the shock until reaching an energy $E_k >> m_e c^2$. This mechanism is also referred to (in the literature) as the diffusive shock acceleration (DSA) mechanism (Pittard and Dougherty, 2006; De Becker, 2007; Palacio et al., 2016). The crucial aspect is the iterative process allowing the particles to move back and forth along the strong shock front, gaining energy in the process. It is thought that magnetic turbulence

and magneto-hydrodynamic waves (Alfvèn waves) in the region of Figure 5.3 labelled as the contact discontinuity (i.e. between the two shocks), essentially scatters the high energy particles (isotropically), allowing them to cross back over the shock front (De Becker, 2007). Whilst magnetic fields are necessary for the production of synchrotron emission, they also facilitate the production of particles at relativistic speeds. Once the particles have reached very high energies, they are able to interact with the magnetic field, spiralling around its field lines emitting synchrotron radiation at a range of radio wavelengths (Blomme, 2011). Whilst this proposed mechanism responsible for accelerating the particles to relativistic speeds cannot be directly proven observationally, several models have employed the DSA mechanism into their hydrodynamical models of PACWBs, which have been successful in reproducing the observed non-thermal emission (Pittard and Dougherty, 2006; Reimer, Pohl, and Reimer, 2006; Palacio et al., 2016).

The remaining 'unknown' is therefore the assumption of a local magnetic field in and around the WCR that now not only promotes the emission of synchrotron radiation but is also a vital ingredient in the DSA (or first order Fermi) mechanism. Magnetic fields within stars, especially early type stars, have often been a large uncertainty within the realms of stellar astrophysics. Much of the theory and knowledge we have of stellar magnetic fields has stemmed from the observations of the Sun, however, how this translates to higher mass stars with large convective envelopes is poorly understood. The recent (and ongoing) Magnetism in Massive Stars (MiMeS) project (Wade et al., 2016) is seeking to clarify our current understanding of the magnetic properties of O and B type stars. In a recent paper by Grunhut et al. (2016), the members of the MiMeS survey report that just $7\pm3\%$ of their sample of 108 individual O-stars (i.e. including the individual members from multi-line systems) are found to have a magnetic signature to a median uncertainty of 50 Gauss (G) on the measurement of the longitudinal magnetic field, $B_l$. Whilst this is a small fraction, for the OB stars that have been found to harbour surface magnetic fields, they have been found to be highly stable and strong, from a few hundred to tens of thousands of G (Petit et al., 2013; Kurapati et al., 2016).

Mageto-hydrodynamical (MHD) simulations (e.g. by Ud-Doula, Owocki, and Townsend 2008; Owocki, Townsend, et al. 2008), in combination with observational constraints have

shown how the stellar winds of such objects interact with their $\vec{B}$ fields, generating stable circumstellar magnetospheres through various wind confinement and channeling processes. Petit et al. (2013) characterised these magnetospheres into dynamical magnetospheres (DM) and centrifugal magnetospheres (CM). Each of which depend upon the degree of magnetic wind confinement as characterised by the Alfvén radius, $R_A$ (i.e. the radius at which the kinetic pressure of the wind becomes equal to the magnetic pressure) and the stellar rotation, characterised by the Kepler co-rotation radius, $R_K$. In the case of a slowly rotating star, $R_A < R_K$ meaning wind plasma confined within the magnetic field is not centrifugally supported and falls back to the stellar surface on a dynamical timescale (Townsend, Owocki, et al., 2006) leading to a DM. For faster stellar rotation, $R_A > R_K$, the wind plasma is centrifugally supported from any infall leading to higher magnetospheric plasma densities (Kurapati et al., 2016). Observationally, Kurapati et al. (2016) conducted a radio survey (with the JVLA and GMRT) of seven and eleven known magnetic O and B stars respectively. Out of four detections (two O stars: known to be in binaries, and two B stars: not known to be in binaries) of their target sample, they deduce that three of these stars are non-thermal radio emitters, which are also predicted to host a CM. They conclude that the detected radio emission of the remaining target (the O-star HD 37742) is mainly of a thermal origin, whilst it is predicted to host a DM (i.e. not a CM). This suggests that non-thermal radio emission is favoured in stars that host a CM, but does not provide evidence as to whether this is also independent of binarity.

In some cases, it has been found that the true extent of the non-thermal emission from PACWB is not detected, since it can suffer from free-free absorption in the stellar winds of either star. The amount of free-free absorption depends on the distance between the two companion stars and also the phase of the binary orbit, since the sightline from the synchrotron emission region to the observer passes through different areas of the wind as each of the stars orbit around their centre of mass. The eccentricity of the orbit must therefore also be taken into consideration, since this can drastically vary the amount of free-free absorption present in the line of sight between the observer and the WCR.

It is clear that the non-thermal radio emission from PACWB is dependent on a variety of parameters and that more observations across a range of frequencies are required in

order to constrain these systems. The recent work of Palacio et al. (2016) outlined a typical radio spectrum of a PACWB, notably that at frequencies between $2 < \nu < 10$ GHz, the non-thermal component of the WCR is dominant and can be determined by the energy of the particles and the local magnetic field strength (also of course the amount of free-free absorption and hence the phase, period, eccentricity of the binary system of that epoch). For frequencies, $\nu > 10$ GHz, the thermal component of the individual stellar winds become dominant (due to its positive spectral index), which is a function of the mass-loss rate, terminal velocity and ionisation conditions of either wind. At frequencies, $\nu < 2$ GHz however, Palacio et al. (2016) suggest that absorption effects such as free-free absorption and synchrotron self-absorption become more prevalent (although synchrotron self-absorption has been shown to be negligible unless the object is very compact Marcote et al. 2015).

## 5.3 Massive star binaries within the COBRaS 21cm radio observations

This section presents the results obtained from an analysis of the massive star binary systems present in the Cyg OB2 association that fall within the field of view of the COBRaS Legacy L-band observations. By first presenting the target sample and their derived flux densities, individual systems are investigated in the following subsections with a comparison to previously published observations.

### 5.3.1 Target sample selection

The initial target selection sample came from the most recent census of massive star binaries within the Cyg OB2 association from Kobulnicky et al. (2014). As part of their radial velocity survey of Cyg OB2, they present orbital solutions for a total of 48 massive OB multiple star systems. The first task was to then limit their sample of 48 objects to only those found within the field of view of the COBRaS L-band pointings. This was accomplished by first finding all the sources within a 400 arc-second radius of the centre

of each of the seven L-band pointings, i.e. representing the entire field of view of the COBRaS 21cm observations. This limited the original sample of 48 objects, to a total of 27, potential PACWB systems within the field of view. Table 5.1 shows the positions of these 27 objects on the sky alongside some of their orbital information as obtained from the various references listed.

Each of the 27 objects (regardless of the system spectral type) listed in Table 5.1 were searched for within the core COBRaS L-band (21cm) observations taken between Apr. 25th and Apr. 27th (from here on Apr. 26th; see Chapter 2 for details on the data reduction process). For each object, a single 512 × 512 pixel image was produced using the AIPS task IMAGR, centred upon their given coordinates as listed in Table 5.1. A natural weighting scheme was used in order to maximise the sensitivity of the resulting images, which were subsequently primary beam corrected using the AIPS task PBCOR. In reaching sensitivities between $\sim 20-24\,\mu$Jy (in the centre of each field), a total number of four objects from the 27 listed in Table 5.1 were detected. This included the previously known PACWB objects Cyg OB2 #8A and Cyg OB2 #9, the quadruple system of Cyg OB2 #5 Kennedy et al. (2010) and the O7.5I-III + O/B binary A11 (or MT91 267) which has no prior detection at radio wavelengths.

For those sources detected in the Apr. 26th data, images and flux densities were also measured from the Apr. 11th epoch. Although the data sets were separated by only 15 days, this timescale could provide a valuable insight into the potential variability of the 21cm flux density, most notably for the short period (P = 21.908 ± 0.040 days) binary, Cyg OB2 #8A. Images of each of the four objects were produced again using IMAGR from the Apr. 11th observations, reaching sensitivities between $\sim 38-44\,\mu$Jy for the centre of each field. Whilst Cyg OB2 #8A, Cyg OB2 #9 and Cyg OB2 #5 were all detected in the Apr. 11th observations, only a $3\sigma$ upper limit could be derived for A11.

The source extraction algorithm for COBRaS (SEAC) was used to derive the flux densities of each of the detected sources across both observation epochs. The program was run using a seed threshold, $\sigma_s = 5$ and a flood threshold, $\sigma_f = 3$ (for further information please see Section 3.1.1). These parameters and hence the inferred flux densities are dependent upon

TABLE 5.1: Positional and orbital information of the 27 potential PACWB systems found within the field of view of the COBRaS L-band observations. The original sample was based upon the census from Kobulnicky et al. (2014). MT91 refers to the ID numbers presented in Massey and Thompson (1991).

| RA (J2000) | DEC (J2000) | MT91 | Other | Spectral Classification | Period (days) | Reference |
|---|---|---|---|---|---|---|
| 308.11047 | 41.32052 | 252 | - | B2III + B1V | 9.558 ± 0.001 | 0,1 |
| 308.13129 | 41.23558 | 267 | A11 | O7.5III-I + O/B | 15.511 ± 0.056 | 2 |
| 308.15427 | 41.38479 | 292 | - | B2V + B-F | 14.8106 ± 0.0011 | 0 |
| 308.15736 | 41.43762 | 295 | - | B1.5V + G-B | 2.4628 ± 0.000 | 0 |
| 308.17872 | 41.3379 | 311 | - | B2V + B3V | 5.752 ± 0.005 | 4 |
| 308.20703 | 41.42679 | 336 | - | B2V + B-M | 2.04087 ± 0.00006 | 0 |
| 308.20842 | 41.39574 | 339 | - | O8V + O-G | 37.84 ± 0.04 | 0 |
| 308.2455 | 41.075 | 372 | - | B0V + B2:V | 2.227 ± 0.001 | 3,5,0 |
| 308.24851 | 41.25407 | 378 | - | B0V + B: | 29.41 ± 0.03 | 0 |
| 308.26217 | 41.29531 | 390 | - | O8V + G-O | 4.6252 ±0.0012 | 0 |
| 308.28696 | 41.22169 | 417B | - | O6V + A-O | 38.0 ± 0.2 | 0 |
| 308.29001 | 41.21682 | 421 | - | O9V + B9V-A0V | 4.161 | 6 |
| 308.29378 | 41.37291 | 429 | - | B0V + (B3V+B6V) | 2.9786 ±0.0001 | 4,6 |
| 308.29474 | 41.25228 | 431 | Cyg 9 | O5-5.5I + O3-4III | 860.0 ± 3.7 | 7 |
| 308.30527 | 41.22465 | 448 | - | O6V + early A-O | 3.1704 ± 0.0003 | 0 |
| 308.31283 | 41.31402 | 465 | Cyg 8A | O6If + O5.5III: | 21.908 ± 0.040 | 8,9 |
| 308.31808 | 41.31716 | 473 | Cyg 8D | O8.5V + O9V + A2V-O | 1687 ± 51 | 0 |
| 308.32513 | 41.36018 | 485 | - | O8V + early B-O | 4066 ± 45 | 0 |
| 308.38185 | 41.36277 | 561 | - | B2V + B-O | 40.09 ± 0.04 | 0 |
| 308.40417 | 41.26981 | 588 | - | B0V + early A-B | 245.1 ± 0.3 | 0 |
| 308.41295 | 41.32385 | 601 | - | B0Iab + B-O | 510.2 ± 0.9 | 0 |
| 308.41583 | 41.38122 | 605 | - | B1V + B1: | 12.27 ± 0.01 | 4 |
| 308.45352 | 41.32795 | 646 | - | B1V + B1.5V | 49.8 $m$0.2 | 0 |
| 308.49803 | 41.29319 | 696 | - | O9.5V + B0V | 1.46918 ± 0.00002 | 4,10 |
| 308.52508 | 41.13583 | 720 | - | B0-B1 + B1-B2 | 4.0677 ± 0.0003 | 2,4 |
| 308.09343 | 41.30527 | - | Cyg 5 | O6.5-7I + O5.5-6I | 6.6 (fixed) | 11, 12 |
| | | | | +late O/early B | 2445.5 ± 109.5 | 13 |
| | | | | + early B | ∼9200[1] yrs | 14, 15 |

References: **0** Kobulnicky et al. (2014), **1** Kiminki, McSwain, and Kobulnicky (2008), **2** Kobulnicky et al. (2012), **3** Kiminki et al. (2009), **4** Kiminki et al. (2012), **5** Woźniak et al. (2004), **6** Pigulski and Kołaczkowski (1998), **7** Blomme et al. (2013), **8** De Becker et al. (2006), **9** Blomme et al. (2010), **10** Rios and DeGioia-Eastwood (2004), **11** Rauw, Vreux, and Bohannan (1999), **12** Hall (1974), **13** Kennedy et al. (2010), **14** Contreras et al. (1997), **15** Linder et al. (2009).

[1]This has been estimated assuming a distance of 1.7 kpc to the object.

the RMS noise which was calculated within a 4 × 4 arcsecond box around the centre of each source. Table 5.2 shows a summary of the source flux densities (or upper limits) for each of the four sources across both observation epochs. The uncertainty on each of the flux densities was derived from a combination of the error upon the measured flux density (a function of the local mean and RMS values) and that of a 10% error associated with the amplitude calibration. For Cyg OB2 #5 there are two fluxes listed for the Apr. 26th epoch as these observations show two individual components associated with the source, one to the NE and the other to the SW. Unfortunately the observations taken on the Apr. 11th were unable to detect the SW component of Cyg OB2 #5, hence a $3\sigma$ upper limit of the flux density is listed. This upper limit, as well as that for the source A11 (within the Apr. 11th observations) were calculated to be 3 × the local RMS of the image.

TABLE 5.2: Positions and fluxes of the four massive star binaries as derived from the COBRaS L-band observations. The positions listed are a weighted position calculated by SEAC and taken from the Apr. 26th observations. Similarly the fluxes listed across both observation epochs were calculated using SEAC with seed and flood threshold levels, $\sigma_s$ and $\sigma_f$ equal to 5.0 and 3.0 respectively.

| Source Name | ID# | RA (J2000) | DEC (J2000) | $F_{21cm}^{26^{th}}$ ($\mu$Jy) | $F_{21cm}^{11^{th}}$ ($\mu$Jy) |
|---|---|---|---|---|---|
| A11 | 8 | 20 32 31.53 | 41 14 08.14 | 219 ± 35 | <98 |
| Cyg 9 | 21 | 20 33 10.73 | 41 15 08.12 | 1307 ± 139 | 791 ± 102 |
| Cyg 8A | 22 | 20 33 15.07 | 41 18 50.42 | 1083 ± 122 | 2817 ± 304 |
| Cyg 5 (NE) | 6b | 20 32 22.48 | 41 18 19.44 | 921 ± 117 | 226 ± 54 |
| Cyg 5 (SW) | 6a | 20 32 22.42 | 41 18 18.88 | 455 ± 74 | <136 |

Each of these four objects are discussed on an individual basis below. Where possible, the 21cm flux densities presented in Table 5.2 are compared to those previously found. Furthermore, the Heliocentric Julian Date (HJD) of each of the sources over both observation epochs was calculated considering the average universal time (UT) over the course of the observations and their coordinates on the sky. This was compared directly to the modified Julian Date (HJD-2450000) corresponding to the time at periastron passage, $T_0$ (i.e. the point in the orbit when the stars are closest to one another) found in the literature. The orbital phase, $\phi$, was calculated by taking the difference between the MJD at $T_0$ and $T_1$, the epoch of each of the COBRaS 21cm observations and dividing by the orbital period of

the system, i.e.

$$\phi = \left(\frac{T_1 - T_0}{P}\right) - \left\lfloor \frac{T_1 - T_0}{P} \right\rfloor \tag{5.1}$$

Note that for previous flux densities found within the literature whose observations epoch, $T_1 < T_0$, then Equation 5.1 was modified to:

$$(1 - \phi) = \left(\frac{T_0 - T_1}{P}\right) - \left\lfloor \frac{T_0 - T_1}{P} \right\rfloor \tag{5.2}$$

The results of these phase calculations for each of the four sources are given in Table 5.3. For the NE component of Cyg OB2 #5 there are currently no values of the orbital phase, $P$ or $T_0$ within the literature meaning values of the phase could not be derived for either of the COBRaS 21cm observation epochs. The errors on the calculated phases have been propagated through considering the error on both the orbital period and $T_0$ for each object. Note that for those systems with relatively short periods in combination with an ephemeris, $T_0 << T_1$, (i.e. A11 and Cyg OB2 #8A) the uncertainty in the phase sampled by these COBRaS observations is very large.

TABLE 5.3: Phase information of the four multiple star systems observed within the COBRaS observations. The phase, $\phi$ sampled by both observation epochs has been calculated using the values of the Period, $P$ and MJD (HJD-2450000) at periastron, $T_0$ included in the table by use of Equation 5.1. The value of $T_1$ used for each of the four sources was 6774.2 and 6758.9 for the Apr. 26th and Apr. 11th epochs respectively. Note, the value of $T_0$ and $e$ stated for the SW component of Cyg 5 refers to that found from the best fit ($s = 0$) model given in Kennedy et al. (2010) (see text for further details).

| Source Name | $P$ (days) | $T_0$ (MJD) | e | $\phi$ Apr. $26^{th}$ | $\phi$ Apr. $11^{th}$ |
|---|---|---|---|---|---|
| A11 | 15.511±0.056[1] | 5575.6±0.8[1] | 0.21±0.07[1] | 0.2790±0.2792 | 0.2879±0.2764 |
| Cyg 9 | 860±3.7[2] | 4020.72±2.55[2] | 0.71±0.016[2] | 0.1937±0.01 | 0.1757±0.01 |
| Cyg 8A | 21.908±0.04[3] | 1807.1±0.9[4] | 0.24±0.04[3] | 0.7254±0.43 | 0.0214±0.43 |
| Cyg 5 (SW) | 2445.5±109.5[5] | 53836±35[5] | 0.69±0.04[5] | 0.2012±0.05 | 0.1950±0.05 |

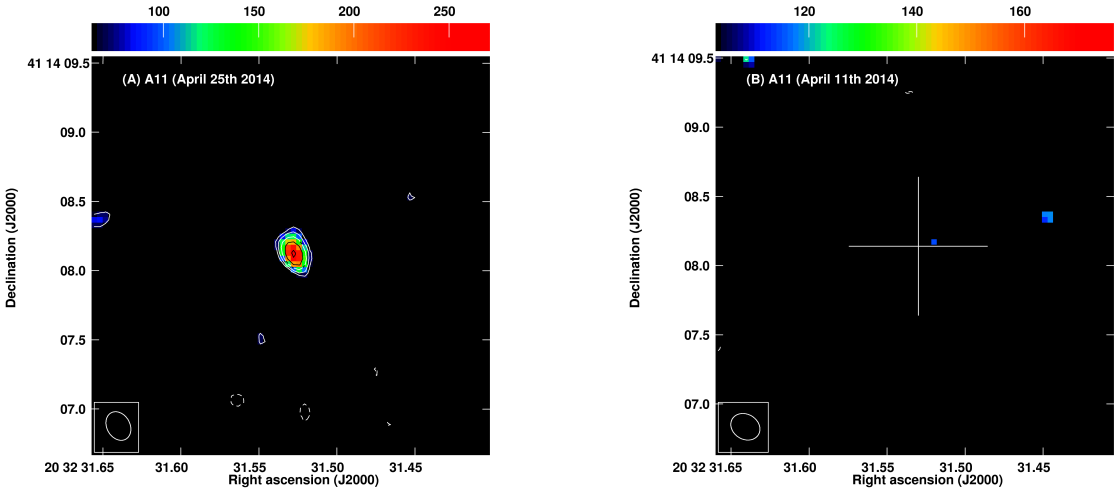
References: **1** Kobulnicky et al. (2012), **2** Blomme et al. (2013),**3** Blomme et al. (2010), **4** De Becker, Rauw, and Manfroid (2004), **5** Kennedy et al. (2010).

### 5.3.2 MT267/A11 (ID: #8)

A11 (in the notation from Comerón et al. 2002) is otherwise referred to as MT267 from the massive stars census of the Cyg OB2 association by Massey and Thompson (1991). Originally classified as a late O-type star (Walborn, 1980) its spectral classification has since been further refined to that of an O7.5Ib-II(f) (Negueruela et al., 2008). It was only recently confirmed as a spectroscopic binary from the radial velocity survey of stars within the Cyg OB2 association by Kobulnicky et al. (2012). Their spectral analysis of the primary star resulted in a spectral classification of O7.5III-I, consistent with that previously found. Moreover, the authors derived an orbital period of $15.511 \pm 0.056$ days and an eccentricity, $e = 0.21 \pm 0.07$ for the system. Whilst unable to derive the spectral type of the secondary, Kobulnicky et al. (2012) suggest its mass ranges from 3.0 to 30.0 $M_{\odot}$ (if adopting a mass of 32 $M_{\odot}$ for the primary) which is dependent on the assumed inclination angle of the system's orbit. Moreover, A11 has been observed to emit in the X-rays, which have additionally been found to be variable, suggesting the presence of a WCR (Rauw, 2011). Assuming this to be the case, the secondary component is likely towards the higher mass end of that previously suggested and implies an orbital inclination angle $i < 13°$ (Kobulnicky et al., 2012).

Further evidence that A11 is indeed a PACWB system is provided in the first ever radio detection of the object from the 21cm COBRaS observations. The system is detected in the Apr. 26th observations with a measured 21cm flux density of $219 \pm 35$ $\mu$Jy (see Figure 5.4a). Considering that no thermal free-free emission from the winds of massive stars was detected within the 21cm COBRaS observations (with the exception of the LBV candidate/binary system Cyg OB2 #12), it is likely that the detected radio emission from A11 is non-thermal in nature and a result of synchrotron emission from a WCR. For the observations taken on Apr. 11th, a $3\sigma$ upper limit of 98 $\mu$Jy was placed on A11, marking a slight variation in the 21cm emission over a $\sim$ 15.5 day period.

Considering the system has a period of $\sim$ 15 days, the two observation epochs are sampling approximately the same orbital phase ($\phi_{26th} = 0.279$ and $\phi_{11th} = 0.288$). For reference, the two flux densities are plotted as a function of orbital phase in Figure 5.5. One would




(A) $26^{th}$ Apr. 2014 (B) $11^{th}$ Apr. 2014

FIGURE 5.4: (A) A11 from the COBRaS Apr. 26th observations, $1\sigma$ rms = 21 $\mu$Jy/beam, colour scale flux range $66 - 270$ $\mu$Jy/beam; (B) a non-detection of A11 from the COBRaS Apr. 11th observations $1\sigma$ rms = 34 $\mu$Jy/beam, colour scale flux range $103 - 176$ $\mu$Jy/beam. The upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms.

expect little variation in the radio emission sampled at the same orbital phase from such a WCR scenario. Note, however, the difference in sensitivity between the two observations epochs and indeed the difference in $u, v$ coverage (see Figures B.1i and B.1j) and hour angle between the two datasets resulting in a different primary beam shape and size. The quoted $3\sigma$ upper limit on the flux density of A11 from the Apr. 11th observations is only a factor of $\sim 2$ less than the flux density as measured from the Apr. 26th observations. However, despite these caveats, this observed variation at almost the same orbital period raises a possible question regarding the period of its orbit. Further observations sampling the entire orbital phase are clearly necessary to derive the radio light curve of A11 to further investigate its likely variable non-thermal emission from a WCR.

### 5.3.3 Cyg OB2 #8A (ID: #22)

Cyg OB2 #8A has been found to be one of the most optically bright objects within the Cyg OB2 association with a visual magnitude $m_V = 9.06$ (De Becker, Rauw, and

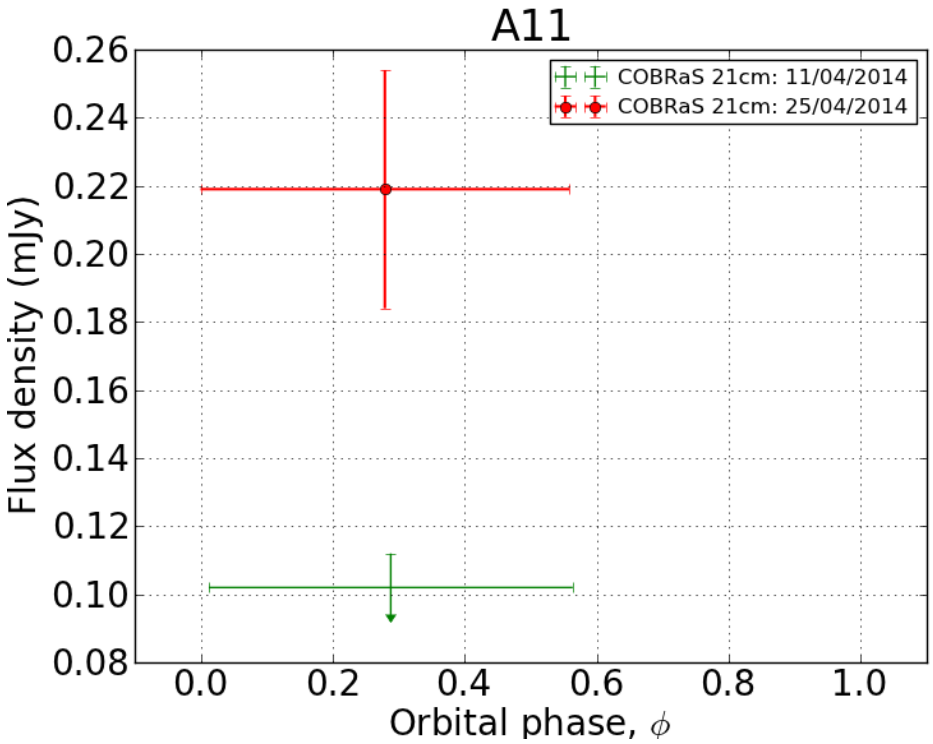
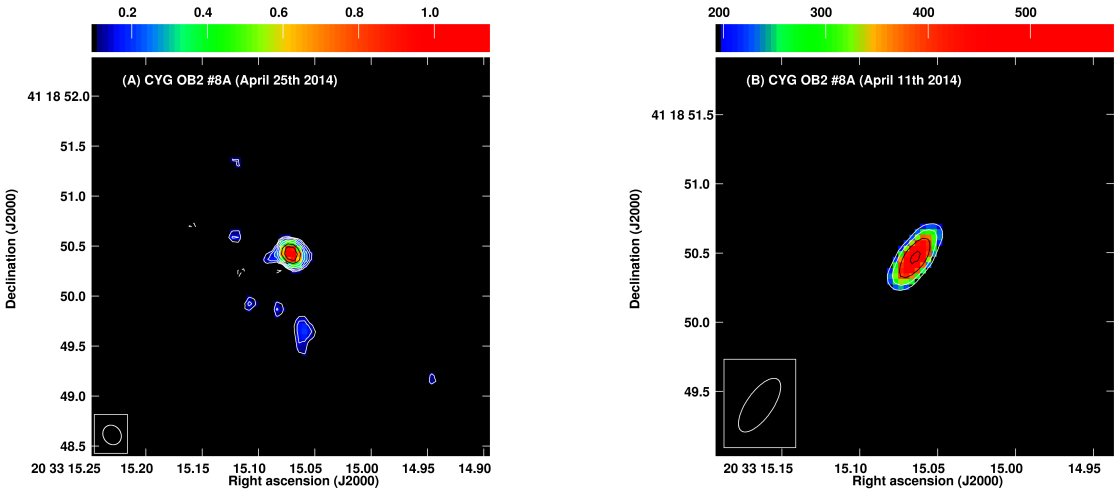



FIGURE 5.5: The two 21cm flux densities plotted as a function of orbital phase, $\phi$ for the O7.5I + early B/ late O PACWB system of A11. The red circle represents the 21cm flux density as measured from the Apr. 26th observations and the green data point represents the $3\sigma$ upper limit provided $\sim$ 15 days previously on Apr. 11th.

Manfroid, 2004). Originally classified as an O5.5I(f) by the photometric and spectroscopic study by Massey and Thompson (1991), it was long suspected to be a binary system due to the presence of strong X-ray emission (see Harnden et al. 1979; Waldron et al. 1998; Kitamoto and Mukai 1996; De Becker et al. 2006) and non-thermal radio emission (see Bieging, Abbott, and Churchwell 1989; Blomme et al. 2010). This was confirmed during a spectroscopic campaign by De Becker, Rauw, and Manfroid (2004) who derived an orbital solution of 21.908 ± 0.040 days for the O6If + O5.5III(f) binary system. Their best fit model to the system's orbit gave an eccentricity, $e = 0.24$ with a mass ratio between the two stars, $q = 1.16 \pm 0.06$.




(A) $26^{th}$ Apr. 2014 (B) $11^{th}$ Apr. 2014

FIGURE 5.6: (A) Cyg OB2 #8A from the COBRaS Apr. 26th observations, $1\sigma$ rms = 34 $\mu$Jy/beam, colour scale flux range $101 - 1140$ $\mu$Jy/beam; (B) a second image of Cyg OB2 #8A from COBRaS Apr. 11th, $1\sigma$ rms = 65 $\mu$Jy/beam, colour scale flux range $195 - 579$ $\mu$Jy/beam. The upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms.

X-ray observations from the XXM-Newton satellite of Cyg OB2 #8A were shown to have a strong phase-locked variability and thus provided evidence that the X-ray emission is likely associated with a WCR between the two components (De Becker et al., 2006). Further evidence for the presence of a WCR was presented by Blomme et al. (2010) who found that the observed radio fluxes taken with the VLA also show a strong phase-locked variability. This was contrary to expectations, since the small period and hence separation between the

two components suggested that at some points throughout the orbit, the WCR (and hence non-thermal radio emission) would be suppressed or hidden by the free-free absorption (due to their stellar winds) within the system. Their model is based upon a synchrotron emission region, orbiting a single star within a stellar wind that includes the effects of free-free absorption, accounts for the geometric extent of the WCR and the acceleration, advection and cooling of the relativistic electrons. Blomme et al. (2010) do not give a formal 'best fit' solution to their model as they could not accurately reproduce the absolute values of the observed flux densities (which they claim are very sensitive to many 'unknowns' within the model, e.g. the magnetic field). They find that their model predicts that the synchrotron emission region extends considerably beyond the apex of the WCR (i.e. further from the regions dominated by free-free absorption), explaining the observed phase-locked variability of the non-thermal radio emission.

Blomme et al. (2010) also present an updated X-ray light curve consistent with that derived from De Becker et al. (2006), constructed using additional observations from the XXM-Newton satellite. The authors find an anti-correlation between the X-ray and radio light curves, likely due to the difference in their respective formation regions. Moreover, the model presented in Blomme et al. (2010) predicts a radio spectral index that differs considerably to the observed value. Since their model is very sensitive to the density of the thermal electrons (i.e. those responsible for the free-free absorption of the synchrotron radiation), they claim that a lower value of the density could reconcile this discrepancy. They state that porosity effects due to optically thick, small-scale wind structures (clumping; see Section 1.2.2) or hydrodynamical effects caused by the WCR could potentially offer such a reduction in the local density. To explore this further, such effects would have to be accounted for in future models of such PACWB systems. None-the-less, it is interesting to note the potential link between the spectral index of the radio emission from PACWB systems and the density of the stellar wind components and therefore wind clumping.

As shown in Figure 5.6, Cyg OB2 #8A has been detected in both epochs associated with the 21cm COBRaS observations with flux densities of $1083 \pm 122$ $\mu$Jy and $2817 \pm 304$ $\mu$Jy for those taken on Apr. 26th and Apr. 11th of 2014, respectively. This marks a 62% decrease in the 21cm flux density over a $\sim$ 15 day period and provides yet further evidence

for the variability in its radio emission. With a period of 21.908 ± 0.040 days, the 15 day separation between the two epochs sample two different values of the orbital phase (see Table 5.3 for details). Figure 5.7 shows the 21cm 'light curve', i.e. the observed 21cm flux density as a function of the orbital phase, $\phi$, for Cyg OB2 #8A for a combination of the values found here (red points), and those found within the literature. The literature values have been taken from Blomme et al. (2010) and represent 20cm observations taken with the VLA from 1982 through to 1999. The blue points represent actual detections where as the green points (in Figure 5.7) are only upper limits.

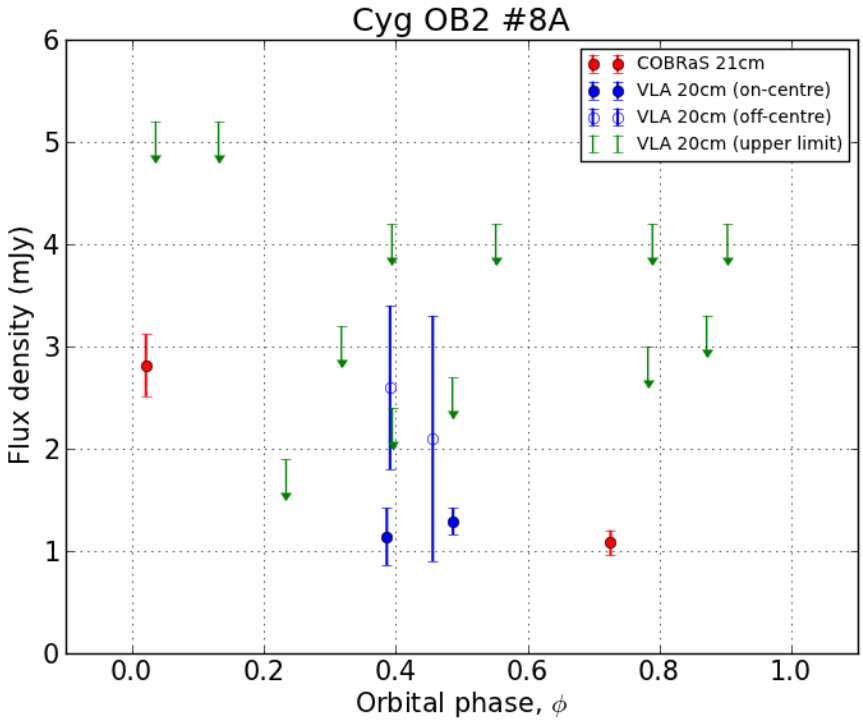



FIGURE 5.7: 21cm flux densities plotted as a function of orbital phase, $\phi$ for the O6If + O5.5III(f) PACWB system Cyg OB2 #8A. The two red circles are those derived from the COBRaS 21cm observations, and the remaining (blue and green) data points have been taken from Blomme et al. (2010). Note: the green data points only represent upper limits; the two blue hollow circles represent measured flux densities from images that are not centred upon Cyg OB2 #8A (see text for further details); the error bars on the phases for the two COBRaS 21cm observations have been omitted from this plot since they are very large (see Table 5.3).

If taking the two data points from the 21cm COBRaS observations alone, one can conclude that the 21cm radio emission from the PACWB systems Cyg OB2 #8A decreases from a value of $\sim$ 2.8 mJy just after the periastron passage, to a value of $\sim$ 1.0 mJy at a phase,

$\phi \approx 0.73$ before increasing again towards periastron. This is confirmed by four observations taken with the VLA (shown by the blue data points in Figure 5.7) and presented in Blomme et al. (2010). These four flux densities cover a phase range between, 0.39 and 0.49, falling between the two phases sampled by the COBRaS 21cm observations. Two of these data points (represented by the filled blue dots and taken from Blomme et al. 2010) are accurate to a similar uncertainty to, and follow the trend set out by, the COBRaS 21cm observations. Whilst the other two (blue circles) data points also follow the trend, they have a far greater associated uncertainty. It is worth noting that these two data points correspond to observations that were not centred upon the Cyg OB2 #8A system but instead on the objects of Cyg OB2 #5 and Cyg OB2 #9 for the data points at phase, $\phi = 0.39$ and at $\phi = 0.46$ respectively.

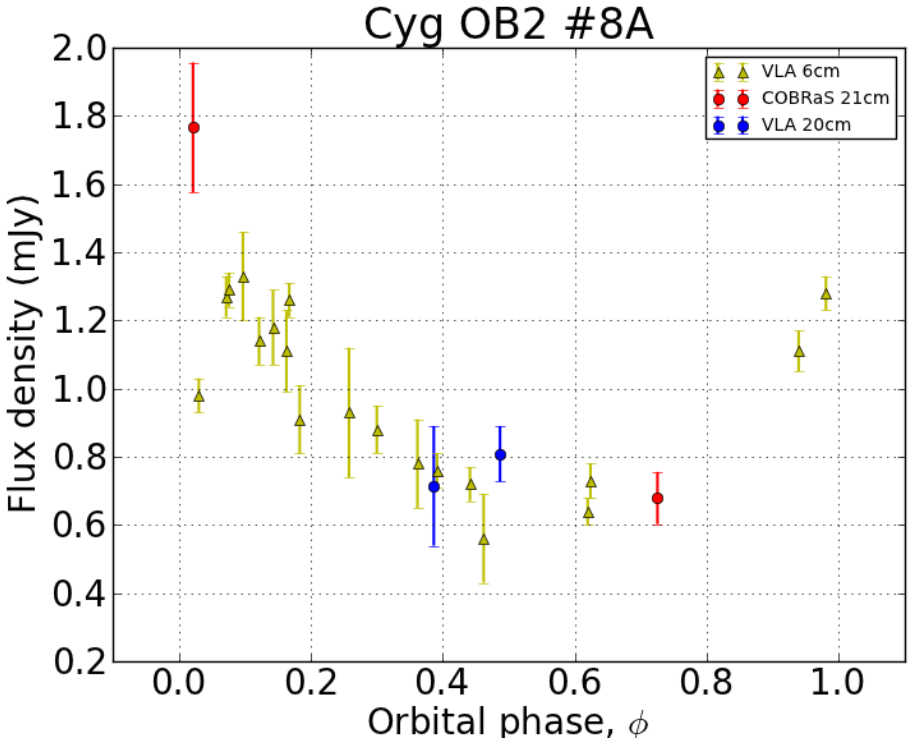


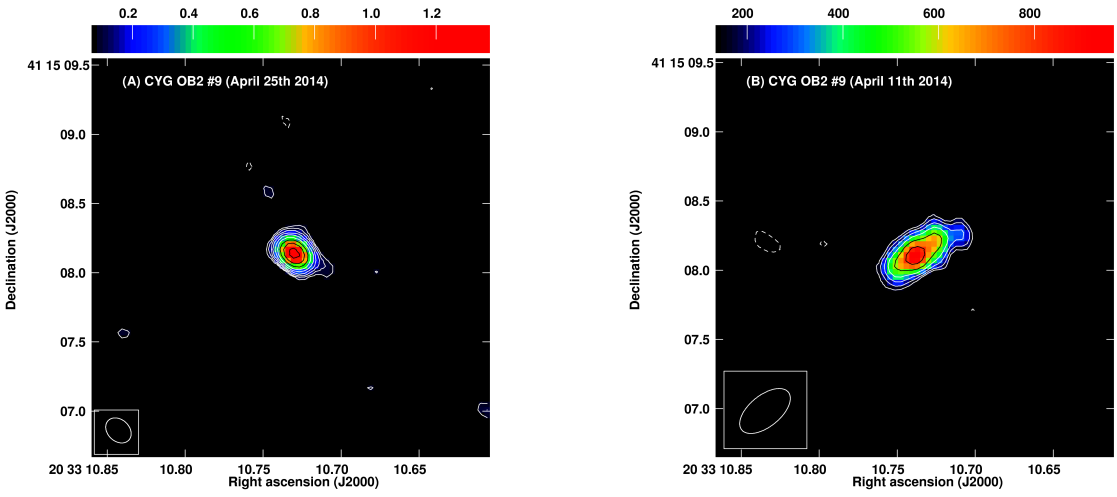

FIGURE 5.8: 21cm and 6cm flux densities plotted as a function of orbital phase, $\phi$ for the O6If + O5.5III(f) PACWB system Cyg OB2 #8A. The two red circles are those derived from the COBRaS 21cm observations and the blue circles and yellow triangles are the 20cm and 6cm flux densities from Blomme et al. (2010). The 20cm and 21cm flux densities have been scaled by a factor of 0.63 in order to normalise them to those representing the 6cm flux densities, this factor was determined from the ratio between the average 6cm and 20cm/21cm flux densities.

Figure 5.8 represents the 6cm flux densities (yellow dots) of Cyg OB2 #8A from Blomme et al. (2010) plotted as a function of phase alongside the 21cm flux densities shown in Figure 5.7. The two 20cm flux densities with the large uncertainty (represented by the blue circles in Figure 5.7) have been omitted from this plot since they were derived from observations that were not centred upon Cyg OB2 #8A. Moreover, each of the 20cm/21cm flux densities have been scaled down by a factor of 0.63 in order for them to co-inside with the 6cm observations. Even though further L-band observations at different orbital phases are required to better sample the variable radio emission at phases between $\sim$ 0.5 and 0.9, the 21cm flux densities derived here are in good agreement with the observed variation of the 6cm radio emission.

### 5.3.4 Cyg OB2 #9 (ID: #21)

Cyg OB2 #9 was first identified as an O5.5If$^{+}$ in the early photometric and spectroscopic observations of the Cyg OB2 association by Morgan, Code, and Whitford (1955) and Walborn (1973). VLA radio observations of the object soon identified it as a non-thermal emitter (Abbott, Bieging, and Churchwell, 1984). However the first real suggestion that it could be a potential binary systems did not surface until the work of Dougherty and Pittard (2006). The authors obtained 8.4 GHz VLBA observations of the system which revealed a conic like shape resembling a bow-shaped WCR. Further evidence of a possible WCR associated with the object came two years later from van Loo et al. (2008) who reported periodic variability in the radio emission on a timescale of $\sim$ 2.35 yrs from the analysis of a large number of archival VLA observations. Direct evidence of its binarity soon came from a spectroscopic analysis of Nazé et al. (2008) who observed a clear doubling of the stellar lines within its spectra. Continued observations determined a preliminary orbital solution for the system, suggesting a long period, eccentric orbit of two stars of nearly equal mass Nazé et al. (2010). Since then, a dedicated observing campaign across X-ray, optical and radio wavelengths has been published as a series of papers titled 'The 2.35 yr itch of Cygnus OB2 #9' (Nazé et al., 2012; Blomme et al., 2013; Parkin et al., 2014).

The optical spectroscopy presented in Nazé et al. (2012) provided a refined orbital solution




(A) $26^{th}$ Apr. 2014 (B) $11^{th}$ Apr. 2014

FIGURE 5.9: (A) Cyg OB2 #9 from the COBRaS Apr. 26th observations, $1\sigma$ rms $= 25$ $\mu$Jy/beam, colour scale flux range $75 - 1368$ $\mu$Jy/beam; (B) a second image of Cyg OB2 #9 from COBRaS Apr. 11th observations, $1\sigma$ rms $= 47$ $\mu$Jy/beam, colour scale flux range $140-960$ $\mu$Jy/beam. The upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms.

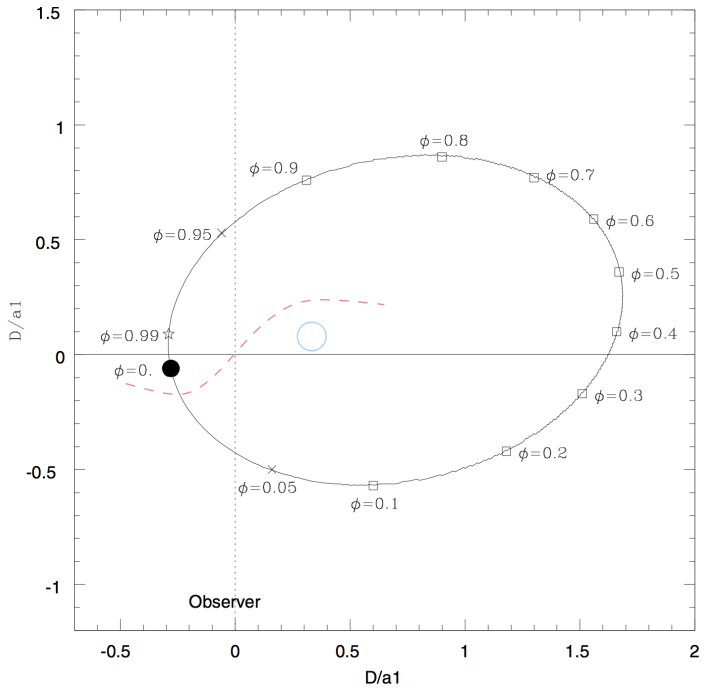



FIGURE 5.10: A schematic illustration of the orbit of the primary component of Cyg OB2 #9 around the centre of mass, derived by and taken from Nazé et al. (2012). The black dot represents the point at periastron, the red dashed curved line represents the possible shape of the WCR and the open blue circle that of the secondary component.

to the system (see Figure 5.10) and further confirmed its spectral classification to be O5-5.5I + O3-4III. Their radio observations at 6cm and 20cm of the system were obtained with the Expanded Very Large Array (EVLA) over the course of 6 months across the periastron passage in 2011. In combination with radio data from van Loo et al. (2008) they infer a period of 860.0 ± 3.7 days for the system and note a steep drop in the radio flux just prior to periastron. This is somewhat expected at periastron where the non-thermal synchrotron emission is subject to a greater amount of free-free absorption from the stellar winds of each component. In paper III (Parkin et al., 2014), a 3D hydrodynamical model of the system is presented and its results are compared to both X-ray and radio observations. In order to reconcile the X-ray fluxes between their model and the observations, they had to scale down the mass-loss rate of the primary and secondary stars by factors of $\sim 7.7 \times 10^{-7}$ and $7.5 \times 10^{-7}$ (i.e. by $\sim 15\%$ and 13%) respectively, from those calculated using standard calibrations as a function of spectral type from Martins, Schaerer, and Hillier (2005) and an $\dot{M}$ recipe from Vink, Koter, and Lamers (2001). Moreover they find an inverse relation between the observed X-ray flux and the binary separation and pre-shock velocity of the stellar winds. Similarly, they find the radio emission to also scale with the separation between the two components and pre-shock velocity, but to positive powers (i.e. in direct contradiction with the X-ray flux). This anti-correlation between the radio and X-ray flux with orbital phase was also found for the case of Cyg OB2 #8A (see Section 5.3.3).

In regards to the 21cm COBRaS observations presented here, Cyg OB2 #9 is detected in both observation epochs with a flux density of 1307 ± 139 $\mu$Jy and 791 ± 102 $\mu$Jy, corresponding to orbital phases of 0.202 and 0.184 for the Apr. 26th and Apr. 11th epochs respectively. These two flux densities are plotted as a function of orbital phase alongside previous observations at 6cm and 20cm from both Blomme et al. (2013) and van Loo et al. (2008) in Figure 5.11. The observations at both 6cm and 20cm show an increase in the radio emission from the time at periastron reaching flux densities of approximately 6 mJy and 8 mJy respectively, until a time shortly before periastron when the radio emission suddenly drops back down to a minimum value. The 21cm COBRaS observations are in good agreement with the variation in the radio flux previously found by both van Loo et al.

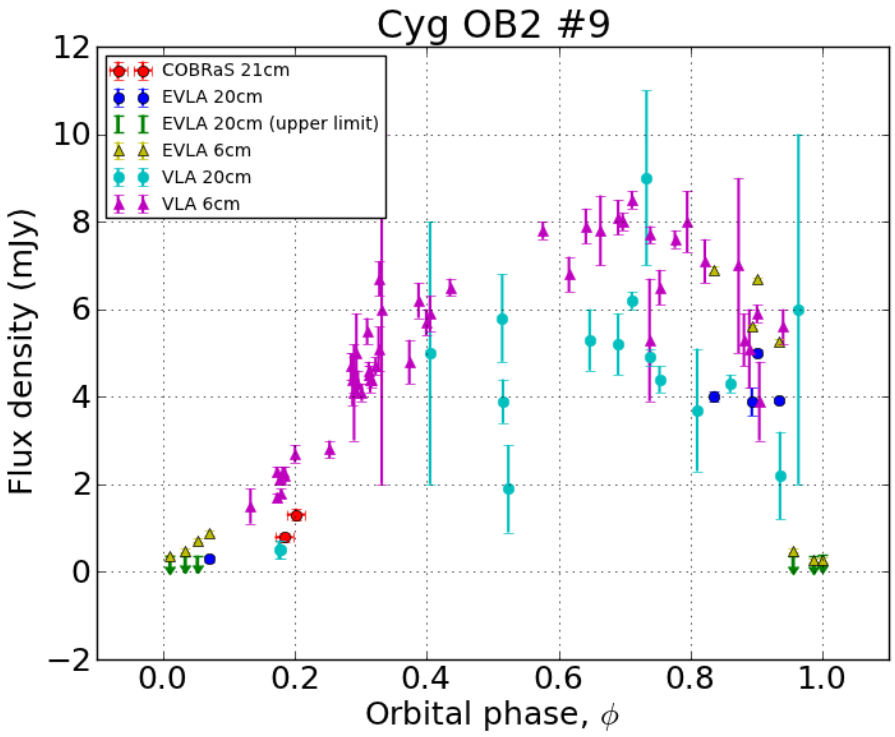



FIGURE 5.11: 21cm and 6cm flux densities plotted as a function of orbital phase, $\phi$ for the O5-5.5I + O3-4III PACWB system Cyg OB2 #9. The two red circles represent the 21cm flux densities from the COBRaS observations, the blue (detected) and green (upper limits) points represent the 20cm flux densities obtained from EVLA observations from Blomme et al. (2013), whilst the yellow triangles mark their EVLA observations at 6cm. The cyan circles and magenta triangles represent the 20cm and 6cm VLA flux densities from van Loo et al. (2008) respectively.

(2008) and Blomme et al. (2013) at both 6cm and 20cm. Similarly to that found with Cyg OB2 #8A, the variation in the radio emission at 6cm and 21cm agree with one another throughout the orbital phase of the system. Both images of Cyg OB2 #9 in Figure 5.9 (i.e. from both of the COBRaS 21cm observation epochs) display a slightly elongated structure, notably to the west of the source in either case. This hints at the possible bow shock like structure expected from a WCR between two massive stars which has previously detected in the case of Cyg OB2 #9 (see Figure 6 of Dougherty and Pittard 2006).

The two red circles representing the COBRaS 21cm data points sample the same oribtal phase, $\phi$, as two of the VLA 6cm observations from van Loo et al. (2008) (to within a difference $\Delta\phi < 0.0015$). Since these are sampling the same orbital phase to a reasonable accuracy, it is possible to determine the spectral index, $\alpha$, on the radio spectrum of Cyg OB2 #9 for these phases, both of which sample the 'low state' of the periodic radio variability. For the phase, $\phi = 0.184542 \pm 0.000664$, the spectral index, $\alpha = 0.85 \pm 0.37$. For the slightly later phase, $\phi = 0.20105 \pm 0.00075$, the spectral index, $\alpha = +0.61 \pm 0.3$. This corresponds to a decreasing thermal spectral index as a function of orbital phase for the radio spectrum of Cyg OB2 #9. This is perhaps expected since at times closer to periastron, when the non-thermal emission is suppressed by an excess amount of free-free absorption along the line of sight, resulting in a radio spectrum dominated by the thermal free-free emission. Unfortunately the COBRaS 21cm observations do not sample the 'high state' in the objects orbital period where non-thermal processes may perhaps dominate the radio spectrum. 21cm flux densities sampling those orbital phases are necessary to determine any possible variability in the spectral index and hence dominant emission processes throughout the orbital period of this object.

### 5.3.5 Cyg OB2 #5 (ID: #6)

Cyg OB2 #5 is believed to be quadruple system (Kennedy et al., 2010). The COBRaS 21cm observations taken on Apr. 26th have succeeded in resolving two individual components of radio emission associated with the system, one to the north east (NE) and the other to the south west (SW) (see Figure 5.13). Figure 5.12 shows a cartoon depicting the four

components associated with the system and highlights the possible WCRs (in red) that could be producing the observed emission components in the COBRaS 21cm observations. Note that each of the four objects are labelled numerically from A to D. This will be the common reference system for Cyg OB2 #5 throughout the remainder of this thesis.

The precise spectral classification of each of the four stars is still very much under-debate. The short period (P = 6.6 days) eclipsing binary at the heart of the system (i.e. stars A and B in Figure 5.12) is thought to harbour two O supergiant stars. This central system was originally classified as an O9+Of binary pair by Hall (1974) using $UBV$ photometric observations. Optical spectroscopic observations of the inner system were later analysed by Bohannan and Conti (1976) who despite deriving a mass ratio between the two inner stars of $q = 4.3$, also concluded that the two components should be of equal brightness. This led to the suggestion that the secondary was in transition from an O-star to a WR object. Leung and Schneider (1978) however, re-analysed the $UBV$ photometric observations from Hall (1974) and concluded that this central system (A+B) was a contact binary with a difference of 2.1 magnitudes between the two objects. Whilst in qualitative agreement with the derived mass-ratio of Bohannan and Conti (1976), this suggestion was contrary to the large spectroscopic signature of the secondary found by the same authors.

Later, variations in the H$\alpha$ line supported the claim that the secondary (star B) could indeed be a WR type object with its observed unusual spectrum resulting from mass transfer from the primary to the secondary Vreux (1985). Further evidence towards this proposition came from an additional set of optical spectra presented in Rauw, Vreux, and Bohannan (1999). The relative strengths of the absorption lines within the spectra gave rise to a spectral type of O6.5-7 and Ofpe/WN9 for the primary (A) and secondary (B) respectively. Additionally, a multi-wavelength approach (optical and X-ray) taken by Linder et al. (2009) also revealed a compact system where bright regions on the secondary star, likely heated by energy transfer from the primary and due to the collision between their two winds, could explain the previous discrepancy between its mass and luminosity ratios.

Evidence for the existence of star C in the Cyg OB2 #5 system has come from variations

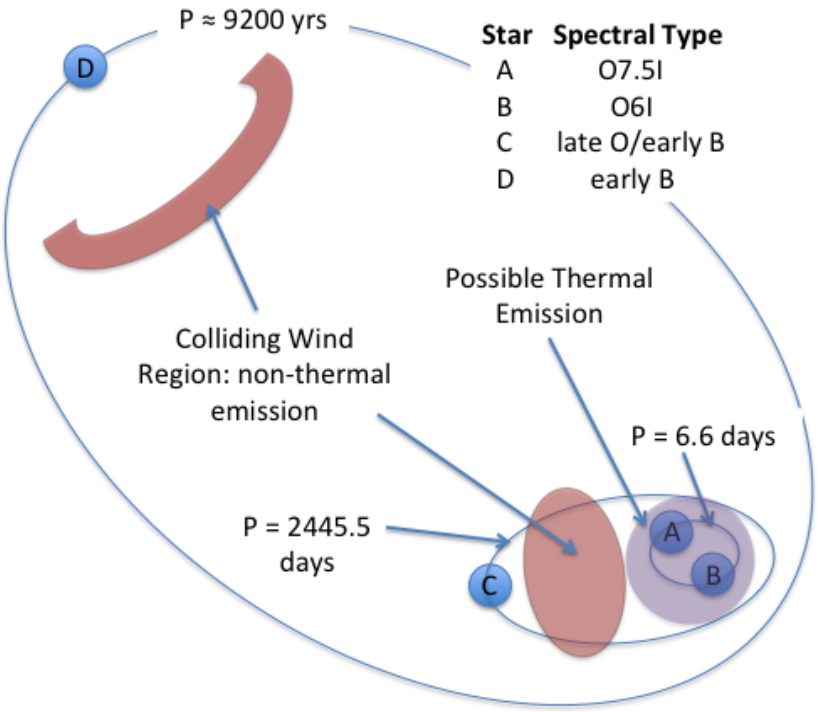
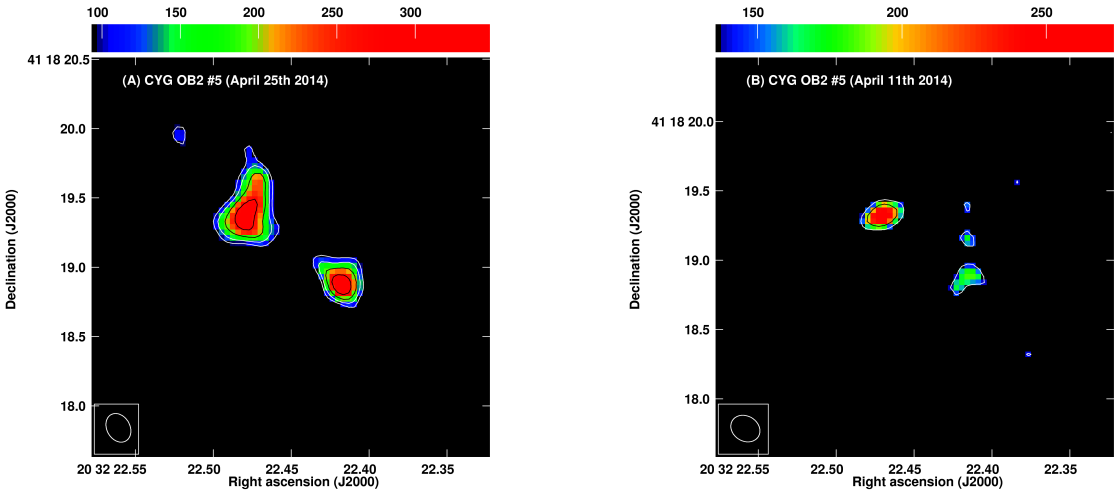



FIGURE 5.12: A cartoon depicting the possible four components that belong to the Cyg OB2 #5 system.




(A) $26^{th}$ Apr. 2014 (B) $11^{th}$ Apr. 2014

FIGURE 5.13: (A) Cyg OB2 #5 from the COBRaS Apr. 26th observations, $1\sigma$ rms = 32 $\mu$Jy/beam, colour scale flux range $95 - 346$ $\mu$Jy/beam; (B) a second image of Cyg OB2 #5 from COBRaS Apr. 11th observations, $1\sigma$ rms = 46 $\mu$Jy/beam, colour scale flux range $137-272$ $\mu$Jy/beam. The upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 $\times$ $3\sigma$ image rms.

in the radio emission associated with the position of the compact binary (i.e. from the SW component in these 21cm COBRaS observations) and the arc-shaped structure observed in the VLBA observations of Dzib et al. (2013). Radio observations obtained with the VLA by Miralles et al. (1994) were the first to report a variation in the radio emission between a high and low state on the timescale of $\sim 7$ years. More recently, all VLA observations of Cyg OB2 #5 over a 20 period were re-examined by Kennedy et al. (2010) who constructed a simple model of a non-thermal radio source orbiting within the stellar wind envelope of the binary. Their derived elliptical orbit (i.e. between the compact binary, A+B and the a third star, C) constrained the variable radio emission to a period of $6.7 \pm 0.3$ years (or $2445.5 \pm 109.5$ days). Using observations obtained with the VLBA, Dzib et al. (2013) find that the 3.6cm flux density of the SW component drops $\sim 7.5$ months prior to the models suggested by Kennedy et al. (2010). Additionally, the authors speculate a possible spectral type for the undetected star C to be between a B0.5 and O8 in order to have a strong enough wind to create the observed WCR when interacting with the combined winds of stars A and B.

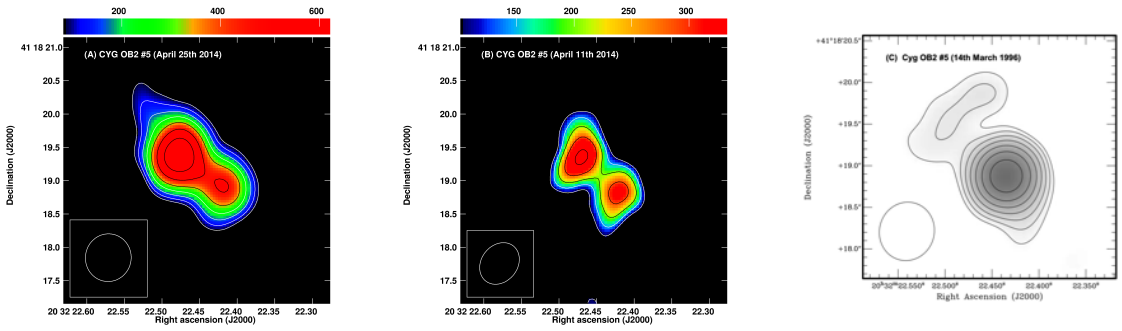
Abbott, Beiging, and Churchwell (1981), Miralles et al. (1994), and Kennedy et al. (2010) have all previously revealed the presence of a non-thermal radio component associated with the Cyg OB2 #5 system, found approximately 0.8 arcseconds to the NE of the position of the compact binary. This is also detected in both epochs of the 21cm observations presented here as shown by the NE component in Figure 5.13a. Kennedy et al. (2010) state that this NE component has shown no signs of variability in the 23 epochs (spanning $\sim 20$ years) where it has been resolved from the (variable) SW component. Moreover, the non-thermal nature of the NE component suggest that it is the result of a WCR between the three stars, A, B and C and a fourth star, D (Contreras et al., 1997; Kennedy et al., 2010). If indeed this is true, then Kennedy et al. (2010) state the required $\dot{M}$ of star D is an order of magnitude larger than that expected for that of an early B-type dwarf, as was suggested to be the spectral classification of star D by Contreras et al. (1997).

#### 5.3.5.1 The NE component of Cyg OB2 #5

The NE component of the Cyg OB2 #5 system, is detected in the 21cm (L-band) COBRaS observations with a derived 21cm flux density of 921 ± 117 $\mu$Jy and 226 ± 54 $\mu$Jy for those taken on Apr. 26th and Apr. 11th respectively. These flux densities were derived from images produced using the AIPS task IMAGR at the full resolution of the array. Unlike previous MERLIN observations at L-band, the 21cm COBRaS observations suggest that the NE component of Cyg OB2 #5 is variable. Note the difference in sensitivity between the resulting images from either epoch. It is possible that a significant portion of the low level flux associated with the source is lost within the noise of the Apr. 11th data, which could potentially account for the observed variation. Moreover, the $u, v$ coverage of the COBRaS 21cm observation epochs fundamentally differ, meaning they are sensitive to different spatial scales on the sky. This is evidenced by plots of their respective (pointing E) $u, v$ coverages shown in the Appendix (see Figures B.1i and B.1j).

The two previous 21cm MERLIN flux densities presented in Kennedy et al. (2010) were extracted from the MERLIN archive and observed approximately 5 weeks apart from one another in 1996. The NE component was initially resolved out by these MERLIN observations due to a lack of low spatial frequencies (i.e. short spacings). As a result, a Gaussian taper of 250 $k\lambda$ at the half width half maximum (HWHM) was necessary to weight up the low spatial frequency data (Kennedy et al., 2010) in order to recover the NE component. It is therefore entirely possible that the 21cm COBRaS observations are also missing flux on large spatial scales. In order to test this hypothesis, a UV taper of 314 k$\lambda$ at 30% of the point spread function (equating to 250 k$\lambda$ at the HWHM) was applied to the COBRaS L-band data for both the Apr. 26th and Apr. 11th epochs. This resulted in a beam size approximately equivalent to that from the MERLIN observations from 1996. The two images of Cyg OB2 #5 from either epoch of the COBRaS observations with the UV taper applied are shown in Figures 5.14a and 5.14b alongside the image produced from the MERLIN observations (Figure 5.14c) taken in 1996 from Kennedy et al. (2010).

The most striking difference between the images of Cyg OB2 #5 as imaged in 2014 with e-MERLIN and in 1996 with MERLIN is the extent of the SW (primary component).




(A) $26^{th}$ Apr. 2014 (B) $11^{th}$ Apr. 2014 (C) $14^{th}$ March 1996

FIGURE 5.14: (A) Cyg OB2 #5 from the COBRaS Apr. 26th observations and imaged with a UV taper of 314 k$\lambda$, 1$\sigma$ rms = 26 $\mu$Jy/beam, colour scale flux range 84 – 619 $\mu$Jy/beam; (B) a second image of Cyg OB2 #5 from COBRaS Apr. 11th observations also imaged with a UV taper of 314 k$\lambda$, 1$\sigma$ rms = 34 $\mu$Jy/beam, colour scale flux range 102 – 332 $\mu$Jy/beam; (C) Cyg OB2 #5 as imaged from a set of MERLIN observations from 14th March 1996 and taken directly from Kennedy et al. (2010), a UV taper of 250 k$\lambda$ was applied at the HWHM (i.e. equivalent to 314 k$\lambda$ at the 30% level as with the images from in A and B), the first contour level is at 3$\sigma$ where $\sigma$ = 0.2mJy, i.e. the images RMS. For (A) and (B) the upper horizontal bar displays the colour scale of each pixel in units of $\mu$Jy/beam, all contour levels are -1, 1, 1.4, 2, 2.8, 4, 5.7, 8, 11.3, 16 × 3$\sigma$ image rms.

The COBRaS observations in 2014 are clearly sampling the SW component during its low state (and indeed the NE component is dominant over the SW), in comparison to the observations from 1996 when the SW component is dominant over the NE component. Due to the lowered resolution in Figures 5.14a and 5.14b SEAC was unable to resolve the NE component from the SW and hence only able to return the combined flux of both components. For the Apr. 26th and Apr. 11th observations this was found to be 1368 ± 147 $\mu$Jy and 746 ± 99 $\mu$Jy respectively. For the Apr. 26th image, this flux density is almost identical to the combined flux densities of the NE and SW components found in the high resolution image shown in Figure 5.13a (i.e. NE+SW = 921 + 455 = 1376 $\mu$Jy).

For the Apr. 11th observations, the low resolution image produced a combined (NE + SW) flux of 746 ± 99 $\mu$Jy, which is significantly higher than the flux of the individual components derived from the high resolution image (NE+SW = 226 + 136 = 362 $\mu$Jy). This suggests that the high resolution image of Cyg OB2 #5 taken from the Apr. 11th dataset (Figure 5.13b) is missing the flux at large spatial scales due to an insufficient $(u, v)$ coverage within the inner baselines. The observed variation of the NE component between

the observations taken with e-MERLIN in 2014 and those with MERLIN in 1996 must therefore be considered as significant, albeit only slightly. These COBRaS observations therefore provide evidence that the 21cm flux density of the NE component of Cyg OB2 #5 has decreased from a value of $\sim$ 1.38 mJy (taking the average value between the two epochs in 1996, see Kennedy et al. 2010) to $\sim$ 0.92 mJy over an 18 year period. Further radio monitoring of the NE component of Cyg OB2 #5 is undoubtedly required.

#### 5.3.5.2 The SW component of Cyg OB2 #5

Unfortunately the SW component, representing the emission from the compact binary and the previously undetected star C has only been detected in the COBRaS 21cm observations taken on Apr. 26th at a flux density of 455 $\pm$ 74 $\mu$Jy. For the Apr. 11th observations, a 3$\sigma$ upper limit of 136 $\mu$Jy was placed on the flux density. This would imply a variation in the flux density by at least a factor of $\sim$ 3 over a relatively short time period ($\sim$ 15 days), note again the significant difference in sensitivity and $u, v$ coverages between the two epochs.

Kennedy et al. (2010) present four different models of the observed variable radio flux of the SW component of Cyg OB2 #5. Within their models, they assume contributions to the radio flux from a thermal component due to the free-free emission within the thermal plasma (i.e. from the stellar winds) surrounding stars A, B and C, and a non-thermal component likely due to a WCR between the wind from stars A+B and C, i.e.

$$S_{obs}(\nu, t) = S_{th}(\nu) + S_{nt}(\nu, t) \tag{5.3}$$

for a given frequency, $\nu$ and a given epoch, $t$. Whilst they assume a constant thermal emission component (which they deduce from their observations at 4.8 GHz during the low emission state), the non-thermal component is dependent on the line-of-sight optical depth. Therefore the non-thermal emission is also dependent on the orbital parameters of the system and the phase of the observations. They approximate that the non-thermal

flux at 4.8 GHz can be represented by a power law:

$$S_{4.8GHz}(t) = S'_{4.8GHz} r^{-s} \tag{5.4}$$

where $S'_{4.8GHz}$ is the non-thermal flux when the separation between the orbiting components (i.e. A+B and C) is equal to the semi-major axis distance, $a$; $r$ is the separation of the two components at time $t$ and $s$ is the power law index. Utilising this relation and solving the equation for the optical depth along the line-of-sight as given by Williams et al. (1990), they determine the observed flux $S_{obs}(\nu, t)$ as a function of the 6.7 year orbital phase, at frequencies of 4.8 and 8.4 GHz, for four values of $s = 0$, 0.5, 1 and 2. The authors claim that each of the four models show good fits to their data at both frequencies and for each of the four emission cycles, observed over a 20 year period with the VLA data. They conclude a best fit solution at a value of $s = 0$ which gives a time of periastron, $T_0 = 3868.5 \pm 35$ (see Table 5.3) and a highly eccentric orbit with $e = 0.7$. The model with $s = 0$ was chosen since it better represents what is expected from a WCR whereby a reduction in the radio flux density (or 'low state') is seen at times close to the periastron, i.e. around a phase $\phi = 0$. Close to the periastron, it is expected that the synchrotron emission from the WCR is subject to a greater amount of free-free absorption, hence causing a net reduction on the radio emission.

TABLE 5.4: Phase information for SW component of Cyg OB2 #5 calculated using the four different models presented in Kennedy et al. (2010). The phase, $\phi$ sampled by both observation epochs has been calculated using a value of the Period, $P = 6.7$ yrs and the different values of the MJD (HJD-2450000) at periastron, $T_0$ shown in the table. The MJD of the Apr. 26th and Apr. 11th observations are 6774.24 and 6758.87 respectively.

| Model $s$ | $S'_{4,8GHz}$ (mJy $a^s$) | $e$ | $T_0$ MJD | Apr. 26th $\phi$ | Apr. 11th $\phi$ |
|---|---|---|---|---|---|
| 0 | $5.3 \pm 0.5$ | $0.69 \pm 0.04$ | $3836 \pm 35$ | 0.2012 | 0.1950 |
| 0.5 | $6.4 \pm 0.6$ | $0.44 \pm 0.04$ | $3516 \pm 34$ | 0.3341 | 0.3258 |
| 1.0 | $7.1 \pm 0.6$ | $0.23 \pm 0.04$ | $3636 \pm 30$ | 0.2831 | 0.2768 |
| 2.0 | $7.6 \pm 0.7$ | $0.11 \pm 0.05$ | $3737 \pm 31$ | 0.24189 | 0.2355 |

The orbital phase sampled by the COBRaS 21cm observations is therefore dependent upon the choice of model from Kennedy et al. (2010). For completeness, Table 5.4 shows

a calculation of the phase sampled by both epochs of the COBRaS 21cm observations for the four different models given in Kennedy et al. (2010). For the $s = 0$, 'best-fit' model (Kennedy et al., 2010), the phases across both COBRaS 21cm observation epochs can be plotted alongside the previous observations obtained at 3.6cm, 6cm and 21cm from the both the VLA and MERLIN (see Tables 1 and 2 of Kennedy et al. 2010). This light curve is shown in Figure 5.15. The main difference between the four models in regard to the flux density measurements is the time at periastron, $T_0$. Varying their parameter $s$ from 0 to 0.5, shifts $T_0$ by approximately a year, whilst the eccentricity sharply decreases with increasing $s$, leading to a value of $e = 0.1$ for $s = 2$. Whilst this alters the physical properties of the system's orbit, the different values of $s$ simply shift the flux densities to different orbital phases.

With reference to Figure 5.15 it is clear that the radio emission at 3.6cm (magenta squares) and that at 6cm (yellow triangles) clearly follow the same variation throughout the orbital period. Similarly, whilst the observed 21cm emission is relatively under-sampled as a function of orbital phase (red, blue and cyan circles), it too shows the same variation as that with the 3.6cm and 6cm emission. However the degree of variation between the measured 21cm flux densities at different orbital phases is much steeper in comparison to that seen at 6cm and 3.6cm. This suggests that the spectral index of the radio emission is changing throughout the orbital period with longer wavelength emission becoming more dominant over shorter wavelengths at a certain phase. This is evidenced by the 21cm COBRaS flux densities, which sit a factor of $\sim 3-4$ lower than the 3.6 and 6cm emission sampled at the same orbital phase ($\phi \sim 0.2$; in Figure 5.15), and additionally (at a phase $\phi \sim 0.5$) by the 20cm emission as measured in 1996 with MERLIN (cyan circles), which are of a similar flux density to both the 3.6cm and 6cm emission. The second COBRaS 21cm flux density (from the Apr. 26th epoch) samples the same orbital phase (to within a phase difference $\Delta\phi = 0.0054$), as one of the 3.6cm and 6cm VLA flux densities. These two VLA flux densities can be used to calculate a spectral index, $\alpha = 0.61$ for that phase, which implies a predicted 21cm flux density of $\sim 1.21$ mJy, i.e. a factor of $\sim 2.6$ higher than the 21cm flux density as measured by COBRaS.

From their VLA measurements at 3.6cm and 6cm, Kennedy et al. (2010) found that the

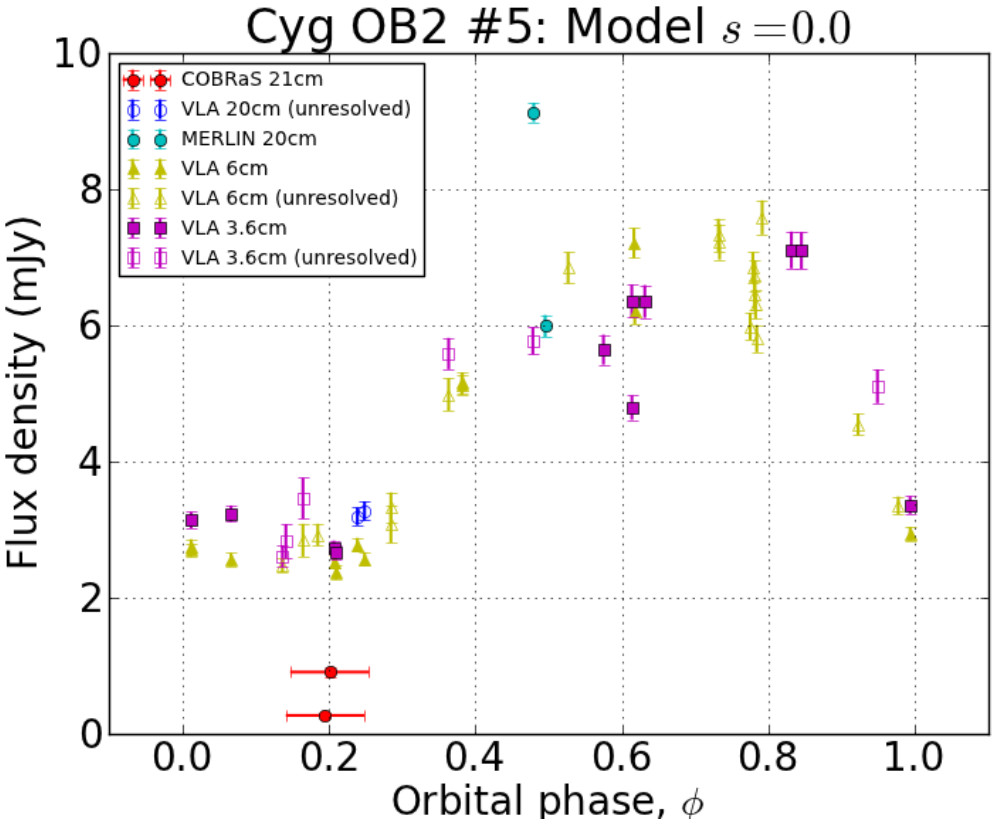



FIGURE 5.15: A light curve of the flux densities at 21cm, 6cm and 3.6cm as a function of orbital phase for the SW component of Cyg OB2 #5. The light curve has been constructed with an orbital period, $P = 2445.5$ days and a time at periastron, $T_0 = 3836$, corresponding so the $s = 0.0$ model of Kennedy et al. (2010). Whilst the two 21cm COBRaS flux densities are shown as red circles, the remaining data points at 20cm (blue and cyan circles), 6cm (yellow triangles) and 3.6cm (magenta squares) have been taken from Tables 1 and 2 of Kennedy et al. (2010) and represent VLA and MERLIN observations taken between 1983 and 2003. Note that all flux densities at 20/21cm and 3.6cm have been multiplied by factors of $\sim 2$ and 0.79 respectively to coincide with the 6cm flux densities. These scaling factors correspond to the ratio between the average flux density of the 20/21cm and 3.6cm to that of the 6cm flux density in either case. An open circle represents a measurement of the flux density of the SW component of Cyg OB #5 having subtracted an average value of the 'stable' NE component at the particular frequency in question, both component were unresolved.

spectral index between the high and low state, varied from $+0.24 \pm 0.01$ to $+0.60 \pm 0.04$. This is suggestive of thermal free-free radiation being the dominant emission mechanism during the low state. This behaviour is confirmed by the low flux densities ($\lesssim 0.5$ mJy) measured by 21cm COBRaS observations which also sample the low state of the orbital phase. Thus when measuring the spectral index of such systems, the orbital phase of the system must also be considered since different emission and absorption features will likely dominate throughout the orbital period. This is especially so for binary systems with more eccentric orbits as is the case with the SW component of Cyg OB2 #5 (if taking the $s$=0 model to be true from Kennedy et al. 2010).

Taking the weighted mean position of both the primary (SW) and secondary (NE) components of the Cyg OB2 #5 system, derived from the Apr. 26th observations it is possible to calculate the separation on the sky between the two radio emission components. The position of the NE component relative to the SW is $+0.882 \pm 0.011$" in RA (east) and $+0.518 \pm 0.010$", in DEC (north) which implies a separation of $1.02 \pm 0.02$". At a distance of 1.4 kpc, this corresponds to a distance of $1428 \pm 21$ AU. Comparatively, the analysis of Kennedy et al. (2010) derived a separation of $0.77 \pm 0.08$" using VLA observations at 8.4 GHz, but found no evidence of any proper motion over the course of 20 years worth of observations. Note however, the difference between the separation of the two components found here and in Kennedy et al. (2010) suggests that the NE and SW components are moving increasingly further away from one another.

In order to better sample the variable radio emission of both the NE and SW components of Cyg OB2 #5, continued observations at radio wavelengths are required. This will also allow for a better handle on the variation of the spectral index (in the case of the SW component) which will give a better indication of the dominant emission and absorption processes as a function of orbital phase. This will provide better constraints on the physical parameters of the system. Furthermore, spectroscopic observations are required in order to detect the presence of (and indeed the spectral type) of stars C and D. This will lead to better estimates of their mass-loss rates and ultimately to tighter constraints on the WCRs within the system.

### 5.3.6 The undetected massive star binary selection

In reference to the original sample of potential PACWB within the field of view of the COBRaS L-band observations (Table 5.1) and considering the sensitivities reached across both observation epochs it is possible to make a statement regarding the type of binary systems expected to be detected. Having detected the known objects of A11, Cyg OB2 #5, 8A and 9, the COBRaS 21cm observations are sensitive to the non-thermal emission produced from the WCR between two or more massive star systems with a spectral type of at least O$_{I-III}$ and that of an early-B, late-O type star, with periods ranging between 15 days and 6.7 years. Moreover, no known systems with such parameters are within Table 5.1 that have not been detected within the COBRaS 21cm observations.

From a morphology perspective, some cases (e.g. the NE component of Cyg OB2 #5) show a bow shaped structure which suggests the presence of a WCR (as was the case with Cyg OB2 #9; Dougherty and Pittard 2006). Additionally, the radio emission from these PACWB systems has been shown to be variable, even on timescales of 15 days (i.e. the time between the two COBRaS 21cm observation epochs). Moreover, the four PACWB have been detected in a range of flux densities between 219 - 2817 $\mu$Jy. With all of these traits considered, it is possible to identify other objects within the COBRaS 21cm observations as potential PACWB systems. This is explored in the next section by only considering objects found to be detected within the COBRaS L-band observations that have also been previously detected at radio wavelengths.

### 5.3.7 Unidentified potential PACWB systems

Considering the large amount of foreground extinction towards the Cyg OB2 association (see Section 1.4.2 for further details), it is possible that unidentified massive star binaries remain within the cluster. Additionally, mid- to long-range period systems ($P > 100$ days) are less likely to be detected spectroscopically. In order to investigate this using the COBRaS 21cm observations, it was necessary to first look at those objects within the COBRaS L-band all source catalogue (CLASC) that have been detected at either radio or X-ray wavelengths (as emission at such wavelengths is expected from PACWBs), that

despite being previously unclassified, had some prior information suggesting that they were at least of Galactic origin and therefore likely within the Cyg OB2 association. If indeed such objects were present within previous catalogues of the region and were detected within CLASC, it is then possible to comment on whether they are likely to be a potential massive star system by considering the following:

1. If they show no or little signs of an extended morphology and/or have a bow-like shape to suggest a WCR.
2. If their measured 21cm flux density was similar to that found for the four detected known PACWB systems.
3. If they were shown to harbour any variation in their 21cm flux densities between the two observation epochs.
4. If they were shown to harbour a negative (non-thermal) spectral index either inferred from previous radio detections or from a combination of previous radio observations and the COBRaS 21cm flux densities.
5. If they have been detected at any other wavelength within previous observations or surveys of the Cyg OB2 association.

The most extensive radio survey of the Cyg OB2 association prior to that performed for the COBRaS project, was conducted using the Westerbork synthesis telescope at 21cm and 86cm (Setia Gunawan et al., 2003). Covering a $2^\circ \times 2^\circ$ region of the sky and centred on Cyg OB2, the survey detected 210 discrete sources (98 at both wavelengths) and a total of 28 resolved sources. A total of 68 of these sources had no previous detection at any wavelength and from a consideration of their spectral index, the authors classify them as either extra-Galactic (44; for those with $\alpha < 0$ and a significant detection at 350 MHz), stellar (10; with $\alpha > 0.3$ and a significant detection at 1400 MHz) and some other kind of Galactic object (14) for the remaining objects. Of these 68 unidentified sources, nine fell within the field of view of the COBRaS L-band observations, from which six were detected within the Apr. 26th observations and five for the Apr. 11th epoch.

These six sources, identified to objects within Setia Gunawan et al. (2003), have been discussed individually within Section 3.4.1. From a discussion of their morphologies, 21cm flux densities between various epochs, and spectral indices (derived where possible), it was concluded that three of these objects are likely extra-Galactic objects (ID #11, #17 and #33) and three are of Galactic origin (IDs #10, #34, and #35). These suggestions were in agreement with the previous classifications given by Setia Gunawan et al. (2003) except for source ID #33 (or SBHW 109), which the previous authors classified as a Galactic HII region. This was based on a detection at 21cm and a non detection at 86cm, however, at the resolution achieved with the e-MERLIN COBRaS observations, two components associated with the object were resolved (see Figure 3.12a), which resemble radio emission from the bi-polar outflows from an Active Galactic Nuclei (AGN). As a result, these observations suggest ID #33 is also of extra-Galactic origin and therefore only source IDs #10, #34 and #35 can be considered as PACWB.

In addition to these three potential massive star binary systems, a further five objects from CLASC, which have been identified with previous X-ray sources, can also be considered to be potential PACWBs. The source IDs of #9 #18 #27 #39 and #41 were all matched to sources from the X-ray catalogue of Wright et al. (2014), which consists of $\sim$ 8000 X-ray point sources. Imaged with the Chandra X-ray observatory, their Legacy survey covers the central 0.5 deg$^2$ of the Cyg OB2 association, similar to the area covered by the COBRaS 21cm observations. Since these objects likely reside within the Cyg OB2 association and have been detected at both radio and X-ray wavelengths, they can be considered as potential massive star binary systems.

This gives a total of eight objects that have been detected within the CLASC and have either a previous radio detection, a previous X-ray detection, or both. Whilst each of these objects were briefly discussed in Section 3.4.1, Table 5.5 gives an overview of their various properties diagnosed from this work. With reference to Appendix B (showing the differences between the $u, v$ coverages of the two COBRaS observations epochs), the flux densities between the two epochs, and their value of $Ss/\theta_{beam}$ (i.e. a measure of how resolved a given object is), each of the eight objects have been classed as either radio variable or not over the 15 days sampled by the two epochs (see Table 5.5). The $u, v$

TABLE 5.5: Information regarding the eight sources potential PACWB objects detected within the CLASC.

| ID | $F_{21cm}^{26^{th}}$ (mJy) | $F_{21cm}^{11^{th}}$ (mJy) | Pointing | $Ss/\theta_{beam}$ $(26^{th})$ | Radio Variable | Morphology | X-ray Source |
|---|---|---|---|---|---|---|---|
| 9 | 0.52 ± 0.07 | <0.15 | B | 2.64 | ✓ | E/B | ✓ |
| 10 | 2.53 ± 0.27 | 1.55 ± 0.18 | E | 8.91 | ✓ | S/E | X |
| 18 | 0.74 ± 0.09 | 0.44 ± 0.08 | G | 3.29 | X | S/E | ✓ |
| 27 | 0.30 ± 0.05 | 0.21 ± 0.05 | F | 2.08 | X | S/E | ✓ |
| 34 | 3.98 ± 0.41 | 3.52 ± 0.38 | F | 5.02 | X | S/E | ✓ |
| 35 | 2.95 ± 0.30 | 1.42 ± 0.19 | C | 5.62 | ✓ | S/E | X |
| 39 | 0.15 ± 0.04 | <0.20 | C | 1.09 | X | E/B | ✓ |
| 41 | 1.30 ± 0.16 | <0.23 | C | 5.75 | ✓ | Ex | ✓ |

Morphologies: (S/E) spherical/elongated; (B) bow shaped; (E) elongated; (Ex) extended.

coverage between the two observations epochs (for a given pointing) and how resolved or extended the object on the sky is, must be taken into account when comparing the two 21cm flux density measurements. Additionally, Table 5.5 contains information regarding the morphology of the eight objects. Note that the same classification scheme as that stated in Section 3.4.2 has been utilised, where a B (bow shaped) type morphology is suggestive of a WCR between two or more massive stars (as was shown to be the base with Cyg OB2 #9; see Dougherty and Pittard 2006).

For each object, their derived properties are all consistent with the four known and detected PACWB systems shown previously. Note however, that for each of the objects also observed to emit in X-rays, their properties are also consistent with the only confirmed YSO observed within CLASC (ID # 38). Additionally, whilst the four objects that are not observed to be radio variable over the 15 day period sampled by the two COBRaS 21cm epochs, a Galactic origin is still likely for these sources, since they are all found within the X-ray point source catalogue of Wright et al. (2014). Of particular note are IDs #39 and #41. Having been detected at both radio and X-ray wavelengths, ID #39 has also been detected in the $J, H, K$ IR bands by the 2MASS survey Cutri et al. (2003). This particular property is also shared by each of the four known PACWB systems of A11, Cyg OB2 #5, Cyg OB2 #8A and Cyg OB2 #9. In regards to ID #41, the 21cm flux density of this object has been shown to increase from a $3\sigma$ upper limit of $< 225$ $\mu$Jy to a detected value of 1297 $\mu$Jy just 15 days later. Such a large variation over a small timescale cannot be reconciled by considering the difference in $u, v$ coverage sampled by the two epochs (which with reference

to Figures B.1e and B.1f is not that significant). This variation must be considered as real and is suggestive of a close period massive star binary system that has been sampled at two opposite ends of its orbital phases. Whilst a flaring event of a pre-main sequence object would also be consistent with this observed 21cm variability, the extended nature of ID #41 (see Figure 3.12d) argues against this conclusion, since such flaring events are likely to be spatially compact.

In order to determine the true nature of these eight potential particle accelerating colliding wind binary systems, follow up radio observations are required to further monitor the radio variability of these objects, over a longer time frame than the separation between the two COBRaS 21cm observation epochs presented here. Furthermore, it is important to note that over one half of the objects detected within CLASC (22 objects) were newly discovered sources with no previous identifications at any wavelength. In Table 3.7, these sources have been classified solely according to their morphologies as observed with the 21cm COBRaS observations. A total of eight of these objects are observed to have a morphology representing a bow shape (B), indicative of a potential colliding wind regions between two or more massive stars.

## 5.4 Discussion and conclusions

This chapter has explored the physical mechanisms behind the production of radio emission from particle accelerating colliding wind binary systems. More specifically, radiation from such objects has been shown to have contributions from both thermal and non-thermal sources. With a positive spectral index, thermal free-free emission from the stellar winds of the individual components dominates the radio spectrum at shorter wavelengths, whilst at longer wavelengths, the non-thermal, negative spectral index of synchrotron radiation becomes the dominant contributor. Section 5.2.2 discussed how the collision between two or more massive star winds leads to the production of relativistic particles via the Fermi mechanism (otherwise known as the DSA mechansim), which in the presence of a magnetic field results in the production of non-thermal, synchrotron emission.

In Chapter 4 it was shown that the 21cm COBRaS Legacy observations from e-MERLIN are in-sensitive (down to a $3\sigma$ flux density limit of $\sim$ 70 $\mu$Jy) to the thermal free-free emission from the stellar winds of single massive stars (with the possible exception of the LBV candidate Cyg OB2 #12). In this Chapter however, it was shown that the COBRaS 21cm observations are sensitive to the non-thermal emission[1] associated with PACWB systems as evidenced by the detection of four systems, all of which have been previously identified as massive star binary systems. These systems include A11, Cyg OB2 #5, 8A and 9, which have been detected with 21cm flux densities ranging between 219 $\mu$Jy and 2817 $\mu$Jy. Their orbital periods range between 15 days (A11) and 6.7 yrs (Cyg OB #5 - SW), the spectral type of the primary is at least an O7.5 star and the secondary (although not always confirmed) is at least that of an early B-type star.

With the exception of A11 (for which only one part of the orbital phase has even been sampled at radio wavelengths), the radio emission of each of these previously identified PACWBs has been shown to vary as a function of orbital phase. For Cyg OB2 #5, 8A and 9, their radio variability has been modelled (see e.g. Kennedy et al. 2010; Blomme et al. 2010; Blomme et al. 2013; Parkin et al. 2014) and shown to be dependent upon a number of physical parameters true to each system. The non-thermal component is dependent on the orbital parameters of the system including the eccentricity and period, the mass-loss rates of either star, the velocity of the pre-shocked gas, the local magnetic field strength, various cooling mechanisms affecting the energy of the relativistic electrons and the optical depth (and hence density) along the line of sight. Such a large parameter space is difficult to model and requires tight observational constraints upon the radio variability of such objects such as those derived here. The thermal component of the radio emission spectrum (as discussed in detail in Chapter 4) is dependent on the stellar wind parameters of the two (or more) stars within the system and should not significantly vary as a function of the system's orbit. As a result, it is a fairly reasonable assumption to treat the thermal free-free emission contributing the radio spectrum as a constant.

[1]Note, it is likely that thermal free-free emission has also contributed to the observed 21cm flux densities of PACWBs.

As shown in the derived light curves for Cyg OB2 #5 (Figure 5.15) and Cyg OB2 #9 (Figure 5.11), their radio spectrum is at a minimum (or low state) at times close to periastron (i.e. when the two components of the system are closest to one another). This implies that the non-thermal emission contributing to the observed radio spectrum is suppressed when the system components are closest to another. Around periastron, the line-of-sight optical depth is likely to increase due the combination of the the stellar winds. As a result, the synchrotron radiation, emitted from the WCR suffers a greater amount of free-free absorption, effectively reducing the amount of radio emission received by an observer. The opposite is seen within the derived light curve of Cyg OB2 #8A (Figure 5.7), whereby a 'radio maximum' is reached around periastron. Cyg OB2 #8A is a much shorter period binary ($\sim$ 21 days) in comparison to both Cyg OB2 #5 and Cyg OB2 #9 and far less eccentric (see Table 5.3 for details). A more circular orbit would suggest that the WCR formed between the two massive star winds is more stable and hence that the radio emission is subject to less variability. This is evidenced by the 6cm flux of Cyg OB2 #8A ranging from $\sim$ 0.6 to 1.3 mJy (see Figure 5.7), whereas for Cyg OB2 #9 $F_{6cm}$ has been shown to vary from $\sim$ 0.16 to 8.1 mJy (see Figure 5.11). The models of Cyg OB2 #8A from Blomme et al. (2010) found that the synchrotron emission regions extend considerably beyond the regions of the stellar wind dominated by free-free absorption. This explains the observed phase locked variability in the non-thermal radio emission from Cyg OB2 #8A and clearly the observed 'radio maximum' at periastron implies the system is subject to a greater production of synchrotron emission relative to the amount of free-free absorption along the line of sight.

The derived light curve of the SW component of Cyg OB2 #5 (Figure 5.15) suggests that the spectral index of the emission is changing throughout the orbital period. This is expected, assuming a constant amount of thermal free-free emission throughout the orbital period but a variation in the amount of free-free absorption that the non-thermal synchrotron emission is subjected to. However the variation in the spectral index may reveal important information regarding the local density of the winds and hence wind clumping. This was first proposed by Blomme et al. (2010) who found that their model of the radio variation of Cyg OB2 #8A could not reproduce the observed radio spectral

index. Since their model was highly sensitive to the density of the thermal electrons (i.e. dependent on the amount of free-free absorption within the stellar winds), they found that a lower electron density was necessary to reconcile this discrepancy. As shown in Chapter 4, wind structures or density in-homogeneities are thought to be present within the winds of hot stars. Therefore by observing the spectral index of PACWB systems and comparing them to models will allow for a further constraint to be made on the density of the thermal electrons and hence if any structure or wind clumping is likely to exist within these hot star winds. Unfortunately, a single frequency observation at 21cm (such as the COBRaS observations presented here) provides no spectral index information. Whilst it is possible to compare the 21cm flux densities to those at other frequencies which sample the same orbital phase, accurate spectral index information cannot be determined unless the same spatial scales have been used when imaging the respective radio visibilities. In order to investigate the potential link between the spectral index of PACWBs and the structure within the winds of their stellar components, the additional C-band observations as part of the COBRaS Legacy project (scheduled to commence in early 2017) are required.

The quadruple star system of Cyg OB2 #5 has been detected in both observation epochs of the COBRaS 21cm dataset. Both of its radio components (SW: primary and NE: secondary) were resolved and likely a result of WCRs between its system members. The primary (SW) component has been found to be variable on a timescale of $6.7 \pm 0.3$ years (Kennedy et al., 2010) and the flux densities of the two COBRaS epochs have been plotted alongside previous flux densities of the SW component at 3.6cm, 6cm and 21cm and sample the low state of the orbital phase (see Figure 5.15). The 6cm radio flux densities of the NE component have previously been found to be stable to within $\sim 22\%$ of the average value, its 3.6cm flux density has varied by $\sim 32\%$ of the average value and no discernible period has been found over the 20 years worth of previous radio monitoring. With only two previous 21cm flux densities of the resolved NE component taken in 1996 (at 1.43 mJy and 1.33 mJy), the 21cm COBRaS observations taken on Apr. 26th show a slight decrease in this value to $\sim 0.92 \pm 0.12$ (some 18 years later). This leads to the conclusion that the NE component of Cyg OB2 #5 is associated with the non-thermal emission from a WCR between the collective winds from the inner three stars A, B and C, and that of a fourth

star D, situated $\sim$ 1 arc-second to the NE of the SW component. Whilst further radio monitoring is required of both components associated with Cyg OB2 #5, the COBRaS C-band will deliver an updated measurement on the 6cm flux density and therefore an updated spectral index of either component.

The L-band COBRaS observations presented here have shown how observations at 21cm wavelengths can be a good indicator of the presence of massive star binary systems. As a result, the COBRaS 21cm flux densities and morphologies for a total of eight objects previously detected at radio and/or X-ray wavelengths are discussed as potential PACWB systems. In order to test this hypothesis further, follow-up observations of each of these objects are required to monitor the potential variability in their radio emission and to derive the spectral index of these objects.

# Chapter 6

# Summary, conclusions and future work

This chapter summarises the work completed in this thesis, drawing particular attention to the main conclusions regarding the technical and scientific aspects investigated throughout the previous chapters. Future work with regards to the ongoing COBRaS Legacy project is discussed, with a focus on continuing and complimenting the science presented in this thesis. Moreover, new observations are proposed in order to build upon the results found from the e-MERLIN L-band Legacy survey in the context of massive star winds.

## 6.1 Thesis summary

The work presented in this thesis has been based on the e-MERLIN 21m (L-band) observations obtained as part of the COBRaS Legacy survey. The underlying scientific goal of this thesis, and indeed the COBRaS project, is to further our knowledge of massive stars, with a particular focus on the properties of their powerful stellar winds and crucially their mass-loss rates. By showing how radio interferometric observations obtained by e-MELRIN can help to improve our understanding of such systems, the focus is drawn to the 21cm L-band observations of the massive stellar association Cyg OB2 which were obtained throughout the spring of 2014. Much of the technical work presented throughout Chapters 2 and 3 deals with the reduction of these data and subsequent image processing and source extraction. This leads to the presentation of the COBRaS L-band all source

catalogue (CLASC). From the results presented in the catalogue, the 21cm radio emission of the known single massive stars within the Cyg OB2 association forms the basis of the work presented in Chapter 4. Whilst constraints are made on both the mass-loss rate and degree of wind clumping within single massive stars, Chapter 5 goes on to explore the incidence of non-thermal emission from the colliding winds of massive multiple star systems.

Chapter 2 presents the techniques and methodology used in the reduction of the entire COBRaS L-band data set. By first summarising the theory behind aperture synthesis imaging with a radio interferometer, the focus is brought onto radio frequency interference (RFI), a prevalent issue across a multitude of radio interferometers and especially so for the e-MERLIN array at the frequencies covered by the L-band ($\sim 1.25 - 1.77$ GHz) receivers. RFI mitigation techniques are discussed and applied to the COBRaS L-band data prior to its calibration. The exact calibration and imaging methods of the data are described in detail, as are the various problems encountered throughout each process. The result of the complete data reduction routine saw that over 50% of the data was lost due to various problems associated with the radio visibilities. These problems ranged from instrumental issues within the receivers, electronics and correlator, an inaccurate recording of the on-source visibilities and RFI, which ultimately lead to a lower limit on the sensitivity of $\sim 22$ $\mu$Jy, compared to the predicted $\sim 10$ $\mu$Jy limit. This highlights the need for both better quality control on future e-MERLIN observations and better instrumental operation.

Chapter 3 discusses the techniques used in deriving the CLASC, the most sensitive 21cm survey of the Cyg OB2 association to date, and provides an initial scientific analysis of the catalogue content. A new source extraction algorithm SEAC that utilises a floodfill algorithm (see e.g. Hancock et al. 2012) is presented and compared to the inbuilt source extraction method, SAD, within AIPS. SEAC is applied to the seven wide field images, which represent the over-lapping pointings from the 21cm COBRaS observations. A thorough methodology was applied to eliminate any potential false detections, resulting in a total number of 41 detected sources above a signal-to-noise ratio, SNR = 3.0. 19 of these sources are matched to previous detections from a large range of catalogues, surveys and observations found within the literature. These include four massive star binary systems;

the LBV candidate Cyg OB2 #12, a class I YSO, at least two possible extragalactic sources (likely AGN) and multiple potential YSOs and massive binary systems. 22 sources have been detected for the first time and have been categorised by their morphological traits. The type of objects detected suggest that these 21cm observations are more sensitive to objects exhibiting non-thermal emission from processes which result in the production of synchrotron emission, in comparison to those whose radio spectrum is dominated by free-free thermal emission.

Chapter 4 then focuses on the free-free thermal emission emitted from the winds of single massive stars. After a discussion of the physical properties of their radiatively driven winds, it was shown how their various observational diagnostics at different wavebands can be used to derive these properties. For a sample of nine, single, massive stars within the field of view of the COBRaS 21cm observations only the blue hypergiant (and LBV candidate) Cyg OB2 #12 was detected[1]. For the remaining stars, strict upper limits on their 21cm flux densities are used to constrain their mass-loss rates, $\dot{M}$, which are compared to values taken from the literature and to those calculated using the theoretical prescription from Vink, Koter, and Lamers (2001). Moreover, a comparison to previous $\dot{M}$ values obtained via $\rho^2$ diagnostics can be used to constrain the amount of small-scale structure (clumping) present within their winds as a function of stellar radii. Whilst the COBRaS 21cm observations suggest that (for at least the hottest stars in our sample) the inner wind regions are more heavily clumped than the outer wind regions, they do not provide any direct evidence for this without a priori knowledge of the star's mass-loss rate. In the advent of highly sensitive mm and radio telescopes such as ALMA and the SKA, observations of the free-free thermal emission from hot star winds will provide better constraints on their mass-loss rates and the radial dependence of $f_{cl}$.

Chapter 5 investigates the non-thermal (synchrotron) emission produced within the wind collision region(s; WCR) between two (or more) massive stars. Considering the large fraction of massive stellar binary systems observed within young massive clusters ($\sim 50\%$; Sana et al. 2012) and more specifically within the Cyg OB2 association ($\sim 55\%$; Kobulnicky

[1]Note that Cyg OB2 #12 is likely part of a binary system, however the measured 21cm flux density from COBRaS has been assumed to have a negligible contribution to any non-thermal emission from any potential WCR associated with the source.

et al. 2014), the importance of massive binaries within the context of stellar clusters and stellar evolution models are discussed. The flux densities of the massive multiple star systems Cyg OB2 #5, Cyg OB2 #8A and Cyg OB2 #9, observed within the COBRaS 21cm observations are presented alongside previous radio observations of each object. For Cyg OB2 #5, these data resolve two clear components to the NE and SW that likely represent two different WCR due to the quadruple body nature of the system. The flux densities of the SW component are plotted against previous radio flux densities as a function of orbital phase and shown to coincide with the observed radio variability. The same is true for Cyg OB2 #8A and Cyg OB2 #9, confirming the radio variability observed in these particle accelerating colliding wind binary (PACWB) systems, due to changes in the line of sight optical depth as the components rotate around their centre of mass. The massive star binary A11 is detected in the radio for the first time and thus also suggests the presence of a WCR between its components. The sensitivity of these COBRaS 21cm observations to the presence of non-thermal emission from PACWB leads to the suggestion that others have been detected within the CLASC. Based upon their morphology, previous identifications, their flux densities as measured between the two observation epochs and in comparison to any previous radio flux, a total of 8 newly detected objects within CLASC are considered to be potential PACWB objects.

## 6.2 The next steps in the COBRaS project

### 6.2.1 L-band data

There is still a significant amount of work that can be done in regard to the COBRaS L-band observations. Firstly the $u, v$ data from each of the seven target field pointings is yet to have been combined together, from which the full field of view covered by the mosaic of the seven pointings can be imaged. Due to the overlapping nature of the pointings it is hoped that this will increase the sensitivity of the regions covered by multiple pointings and thus increase the total number of detections. If we assume that for the area covered

by two pointings, the relative integration time is doubled, this would result in an increase in the sensitivity by a factor of $\sim \sqrt{2}$ for those regions.

An additional number of sources may also be detected if the observations were imaged at the field of view corresponding to that of the five 25m antennas within the e-MERLIN array. This would significantly decrease the sensitivity of the resulting images due to the loss of the Lovell antenna from the array (i.e. the largest collecting area of the array), but increase the field of view of each pointing by a factor of $\sim 3$. As a result, any bright sources (on $\sim$ mJy scales) would likely be detected outside the field of view of that determined by the Lovell antenna.

Alongside the two observation epochs that contribute to the COBRaS L-band Legacy survey made on 2014 Apr. 26th and 11th, there were also some observations taken on the 24th of January 2014 that have not yet been completely calibrated. These observations only account for a small amount of time spent on source ($\sim 0.5$ hrs per pointing, prior to any flagging of RFI or corrupt data), but for the brightest sources within the field they will provide an additional flux density measurement at a different epoch. This will determine if any of the sources appear to be radio variable over the timescale of $\sim 4$ months which will not only help to identify any potential massive star binary systems, but will also add further information to the derived light curves of the known PACWB systems Cyg OB2 #5, Cyg OB2 #8A and Cyg OB2 #9.

In order to make a fairer comparison between the measured flux densities of resolved sources within the two observation epochs (or three for those that may possibly be detected in the 24th of January epoch), a $u, v$ Gaussian taper can be applied when imaging each source individually. A $u, v$ taper will weight the $u, v$ visibilities by a Gaussian of a specified FWHM in $k\lambda$, increasing the detectability of an extended source at the cost of resolution. For data sets whose $u, v$ coverages differ significantly from one another such that one is potentially missing flux on large (extended) spatial scales, a $u, v$ taper can be used to bring both data sets to a similar resolution, and thus making them sensitive to similar spatial scales. Note however, that this process cannot fundamentally change the observational restrictions imposed by the $u, v$ coverage of the observations. A completely fair comparison between

two observation epochs of a given interferometer can only be made if their visibilities occupy the same region of the $u, v$ plane. Never-the-less a $u, v$ taper will increase the reliability of any variation observed between the two observation epochs for extended (resolved) sources. Regardless of these additional objectives for the L-band data sets, the complete COBRaS L-band all source catalogue must also be made publicly available in a virtual observatory (VO) table format.

### 6.2.2 C-band data

As previously stated, the COBRaS C-band (6cm) observations are due to commence in the beginning of 2017. The C-band data set will cover 252 hrs worth of observing time over 42 pointings, contributing to a significant portion ($\sim 5/6$) of the total awarded allocation time of the COBRaS project. Much of the work presented in this thesis will aid the subsequent data reduction, imaging, source extraction and cross-correlation of the C-band data set. It is expected that the C-band data will only be subject to small amounts RFI in comparison to the L-band data set. This is due to the protected frequency space associated across the 2 GHz bandwidth of the C-band receivers installed on the e-MERLIN antennas. As a result, large portions of the calibration routine will likely need to be automated by using the various pipelines available (e.g. see Peck 2014 and the e-MERLIN pipeline made publicly available online[2]). Moreover, due to the sheer volume of data, it is expected that the observations will be made over a number of different epochs, naturally splitting the data into chunks that can be processed individually. Each epoch must be treated in a consistent manner such that they can be combined together at the imaging stage to achieve the required sensitivity limit of $\sim 3\mu$Jy, and thus the calibration will benefit from using the automated pipelines available.

The source extraction algorithm for COBRaS (SEAC) will be the primary source extraction method for the C-band data. This will ensure a good consistency between the source flux densities as derived between the L- and C-band observations, giving rise to more

[2]http://www.e-merlin.ac.uk/observe/pipeline/

reliable spectral index information. Additionally, the same source identification and cross-correlation procedure used in Chapter 3 for the 21cm sources, will be used on the objects found in the C-band data, including the matching of each source to the same catalogues as those described in Section 3.2. For the un-identified objects within CLASC that are also detected within the C-band observations, the measured 6cm flux densities in combination with the 21cm flux densities will provide the spectral index of these objects for the first time. This will help to better classify these objects by finding whether their radio spectrum is dominated by thermal or non-thermal emission. Additionally, if the C-band observations consist of more than one epoch, an investigation of their 6cm radio variability can also be made, helping to further categorise and classify these previously un-identified objects.

In reaching a $1\sigma$ 6cm flux density limit of $\sim 3\mu$Jy, it is expected that the C-band observations will be far more sensitive to the thermal (i.e. positive spectral index) free-free emission emitted from the winds of hot stars, in comparison to the L-band observations. The 6cm flux densities of numerous massive single stars, including non-supergiant O-type stars for the first time, will not only provide tight constraints on their mass-loss rates, but also on the clumping factor $f_{cl}$ in the radio emitting region of their stellar winds. This will allow for a better understanding of the structure in the outer wind regions as a function of spectral type. Moreover, in combination with other multi-wavelength observations, these data will determine the radial stratification of wind clumping, and present new insights into the current disparity between various mass-loss diagnostics.

In regard to the non-thermal emission from PACWB systems (presented throughout Chapter 5 of this thesis), the C-band data, if consisting of a number of different observation epochs, will provide further radio monitoring of the variability of these objects. For any of the systems where the 6cm observations sample the same orbital phase as those sampled in the 21cm observations, the spectral index of that orbital phase can be determined. As discussed for the case of Cyg OB2 #8A by Blomme et al. (2010), when modelling their radio variability, the observed spectral index of these systems can provide a constraint on the density of thermal electrons within the system. The spectral index information, can therefore potentially provide a valuable insight into the density structures of the winds of these stars and hence the degree of clumping within their stellar winds.

## 6.3 Future proposals

Whilst the scheduled COBRaS 6cm observations will vastly improve upon the results obtained from this L-band Legacy survey, it is clear that follow-up observations will be required to further investigate a number of key results. Firstly, in order to sample the free-free thermal emission throughout the different regions in the wind of massive stars, observations at other wavelengths are required. Millimetre (mm) observations for example, will sample the intermediate regions of the O-star winds ($R \sim 10 - 50\ R_\star$). The radial distance sampled will increase with the wavelength of the observations, thus for stars of a given mass-loss rate (regardless of the exact value of $\dot{M}$), multi-wavelength observations from the mm through to the radio will track how the clumping factor $f_{cl}$ changes as a function of radial distance. For a thorough treatment of a given O- or B-type star, these multi-wavelength observations from the mm to the radio would also benefit from observations of the near infrared (IR) Br$\alpha$ line. The HI Br$\alpha$ line (due to the transition between 5th to the 4th energy level of the hydrogen atom) is considered as one of the most promising mass-loss diagnostics for stars with weak winds (Sonneborn and Bouret, 2011). Moreover, observations of their IR continuum can be used to probe the regions of the wind between $\sim 3-10\ R_\star$. Such observations will be key to investigating how the density inhomogeneities (or clumps) evolve as the wind material expands radially outwards from a given star. In order for these observations to be feasible however, a sample of nearby O and B supergiants must be carefully selected such that each instrument observing at a given wavelength is capable of observing them in terms of their sensitivity and position with respect to the coordinates of the sources on the sky.

The L-band observations of Cyg OB2 #12 show that its radio flux was found to be variable over a period of just 15 days. Note however, that the object is resolved in both epochs and the difference between the $u, v$ coverages of either epoch may play a role in this variation. Moreover, due to the recent findings of both Caballero-Nieves et al. (2014) and Maryeva et al. (2016), second and third counterparts have been found for this object. This suggests that at least a portion of the detected 21cm flux is potentially non-thermal and associated with a WCR between the three components of the system. Further radio monitoring is

therefore required of this object in order to better sample any potential radio variability it may possess. Such an investigation would benefit from further multi-epoch e-MERLIN observations at L-band, and at additional wavelengths such as K-band (1.36cm) from which the spectral index could be calculated to see if this too varied. A varying spectral index would be suggestive of a variable non-thermal component associated with the system and give strong evidence for the presence of a WCR within the system.

The radio variability of the SW component of Cyg OB2 #5 is still uncertain, with recent VLBA observations (Dzib et al., 2013) suggesting a decrease in its 3.6cm flux density a full 7.5 months earlier than the models predicted in Kennedy et al. (2010). The 21cm observations presented here, whilst consistent with the 'low radio' state for the sampled orbital phase, are much lower than perhaps expected. In order to better constrain the 21cm radio variability of this object, follow up e-MERIN L-band observations are required to sample the entire orbital period of this object (6.7 yrs). In comparison to the 3.6cm and 6cm observations, this will give a clear indication of how the spectral index of the object changes throughout the orbital period, providing better constraints to accurately model the radio variability. These follow up observations will also sample the 'stable' NE component and monitor any potential long-term variability of this potential WCR. Finally, the nature of the third and fourth (stars C and D) components are still unresolved and dedicated spectroscopic and photometric observations are required to determine their spectral types.

The L-band e-MERLIN Legacy survey of the Cygnus OB2 association has imaged all sources down to a $1\sigma$ 21cm flux density limit between $\sim 22 - 50$ $\mu$Jy depending on their position in the primary beam covered by each of the seven target field pointings. These observations are sensitive to spatial scales down to $\sim 180$ mas and cover an area of $\sim 0.48$ deg$^2$ on the sky. The positions and integrated fluxes of the 41 objects will be of long standing value and in combination with scheduled COBRaS C-band observations, will provide the most extensive radio survey of one of the largest collections of OB stars to date. One of the key results derived from 21cm COBRaS observations was the large number of variable sources as measured between the two observation epochs. In particular, five objects were found to be highly variable, with their 21cm flux density increasing by over a factor $\sim 4$ from the Apr. 11th to the Apr. 26th epochs. Moreover, each of these objects

were found to be heavily resolved in the observations taken on Apr. 26th and only one was detected in the Apr. 11th epoch. Follow up radio observations are clearly required in order investigate these puzzling objects.

The primary intention of this thesis was to further investigate the radio emission from massive ($M \gtrsim 8$ M$_\odot$) stars in the Cygnus OB association in order to obtain a better understanding of their mass-loss rates and binary fraction. Both are fundamental in characterising the stellar evolution of massive stars and are thus hugely consequential to the relative fraction of Wolf-Rayet (WR), red supergiants (RSG), luminous blue variables (LBVs), neutron stars, black holes, gamma ray bursts and supernovae found within our Universe. The next generation of highly sensitive radio and mm telescopes such as the Square Kilometre Array (SKA) and the Atacama Large Millimetre Array (ALMA) will be revolutionary to massive star research. Building on the findings from the e-MERLIN L-band survey of Cyg OB2, future detailed surveys with both ALMA and the Australian Telescope Compact Array (ATCA) will derive the radio properties of massive stars within several other young stellar clusters such as Danks 1 & 2 and the arches cluster. Additionally, large scale surveys with future instruments such as the SKA will bring huge improvements in observational capabilities, reaching the sensitivities required to probe the winds of less evolved massive stars, allowing for reliable mass-loss rates to be determined across a wide range of spectral types. In combination with future instruments such as the James Webb Space Telescope (JWST), such observations will allow for a comparison between different mass-loss diagnostics, leading to a better understanding of massive star wind structures (e.g. clumping) as a function of stellar radii. In addition, the up-coming Gaia mission promises to deliver proper motions, radial velocities, distances and the photometry of $\sim 1$ billions stars. Such information will be highly valued to the on-going research into both massive stars and massive stellar clusters, allowing for vital quantitative tests of massive star evolution whilst improving the number statistics of various evolutionary phases.

# Appendix A

# AIPS functions and nomenclature

The Astronomical Image Processing System (AIPS) has been the primary tool used in the complete reduction of the COBRaS L-band data set. The system has many inbuilt functions or 'tasks' that enable the user to manipulate their data. Here I give a brief definition the main AIPS tasks used throughout the completion of the work presented in this thesis. For further details, please refer to the online AIPS 'cookbook' which can be found at the following *url*: http://www.aips.nrao.edu/CookHTML/CookBook.html.

TABLE A.1: The definitions of the shorthand AIPS table/file names that can be associated with a given data file.

| | |
|---|---|
| AN | Antenna Table |
| BP | Bandpass Table |
| CC | Clean Component |
| CG | Clean Beam Table |
| CL | Calibration Table |
| FG | Flag Table |
| FQ | Frequency Table |
| HI | History File |
| NX | Index File |
| PL | Plot File |
| SN | Solution Table |
| SU | Source Table |

**BPASS:** Calculates the bandpass response functions of each of the antennas, storing them into an AIPS bandpass (BP) table.

**CALIB:** Computes a given calibration routine, storing the solutions (which can later be applied to the data) in an AIPS solution (SN) table.

**CCMRG:** Sorts the contents of a given clean component (CC) table to bring all components sharing the same position together.

**CLCAL:** Applies a given set of solutions stored in one (or more) SN table(s) to those in a given calibration (CL) table, writing them to a second CL table.

**CLIP:** Flags all visibility amplitudes within a given data set that are above a given threshold value.

**CLCOR:** Allows for various corrections to be made to a given CL table.

**FITLD:** Loads both $u, v$ data and maps (stored as FITS files) from a hard disk into AIPS.

**FLATN:** Interpolates (combines) a set of images (e.g. multiple fields produced in IMAGR) into a single image of standard geometry.

**FRING:** Computes both the group delay and phase rate calibration, storing the solutions in an AIPS SN table.

**GETJY:** Determines flux densities of a given source(s) from the CALIB solutions stored in a particular SN table, writing the results in the source (SU) table associated with the data file.

**IBLED:** Allows the user to edit (flag) their visibilities interactively which are displayed as a function of time (averaged in frequency) on the AIPS TV.

**IMAGR:** The sole imaging/clean task within AIPS, offering the full range of weighting, gridding and clean component correction options.

**IMEAN:** Calculates the mean and rms brightness for a given portion of an image.

**INDXR:** Creates an index (NX) table (and a CL table if none previous existed) and index's the $u, v$ data file.

**PBCOR:** Applies a given primary beam correction (i.e. divide by the primary beam gain factor) to a specified image.

**POSSM:** Creates a plot of the $u, v$ data in frequency space (i.e. averaged in time), either displayed on the AIPS TV or stored in a plot (PL) file.

**PRTCC:** Prints the number of clean components computed from IMAGR, associated with a given image.

**QUACK:** Flags selected portions of a $u, v$ data file, according to a set amount of time from the beginning or end of each scan.

**REWAY:** Modifies the $u, v$ data weights of a given data file according to the uncertainty (rms) calculated using multiple spectral channels.

**SETJY:** Enters source information into the (SU) table associated with a given data file.

**SNEDT:** Edits a given calibration table interactively using the AIPS TV.

**SNPLT:** Plots a given calibration table associated with the data file either displayed on the AIPS TV or saved in a plot (PL) file.

**SOUSP:** Fits the fluxes for a given source to determine the spectral index and spectral index curvature.

**SPFLG:** Grids a selected portion of the $u, v$ data file to form a 3-dimensional image across time and frequency space, allowing the user to select and flag particular sections of the plot.

**UVFIX:** Computes the $u, v, w$ values for a given data file with the option to shift the phase centre to a given location on the sky.

**UVFLG:** Flags a selected portion of the $u, v$ data file, storing the flags in a flag (FG) table.

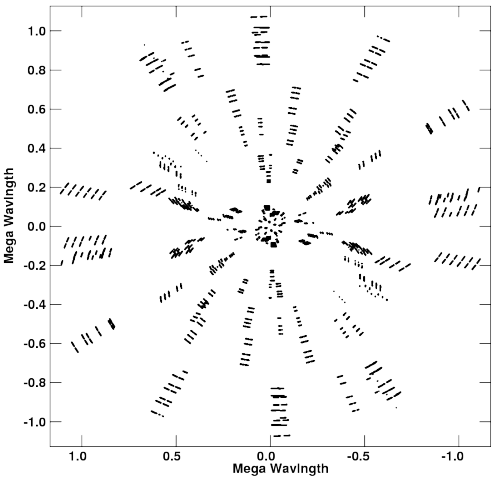
**UVSRT:** Sorts the $u, v$ visibilities into a specified order.

**VPLOT:** Creates a plot of the data in time space (i.e. averaged in frequency) either displayed on the AIPS TV or saved as a plot (PL) file.

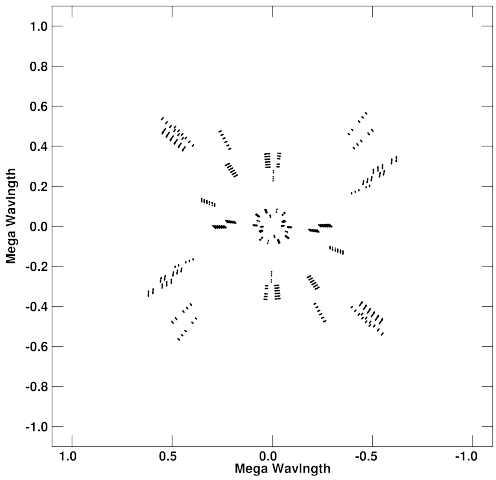
**WTMOD:** Modifies the $u, v$ data weights according to those set by the user.

# Appendix B

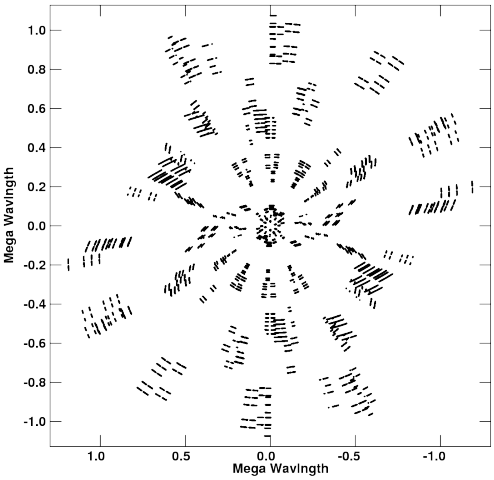
# The $u, v$ coverages of the seven COBRaS 21cm pointings.



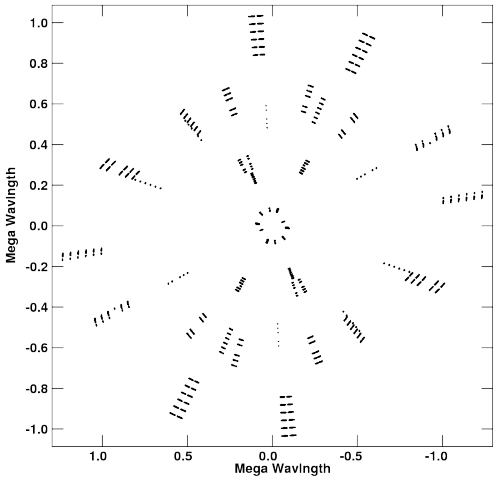

(A) **Point A** ($26^{th}$)



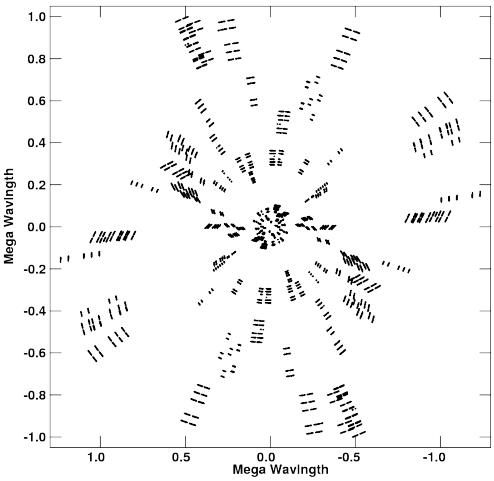

(B) **Point A** ($11^{th}$)



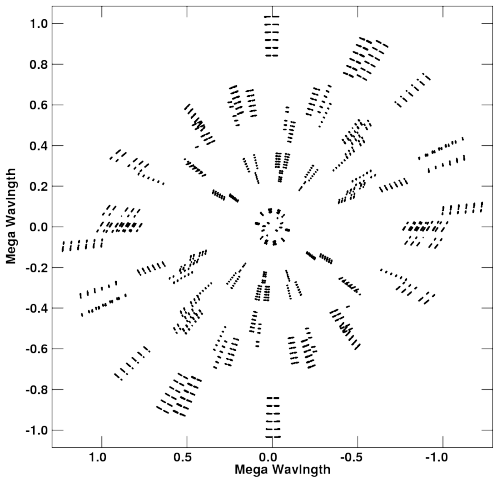

(C) **Point B** ($26^{th}$)




(D) **Point B** ($11^{th}$)

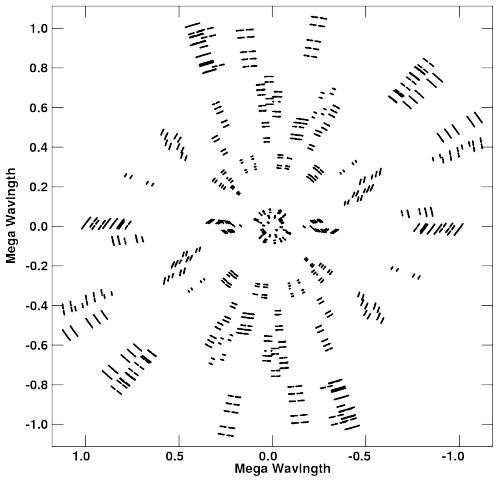



(E) **Point C** ($26^{th}$)

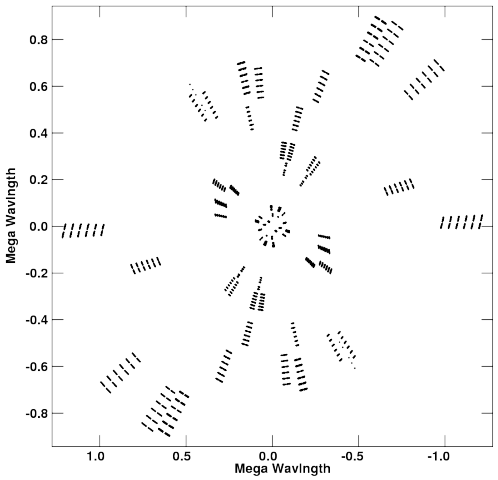



(F) **Point C** ($11^{th}$)

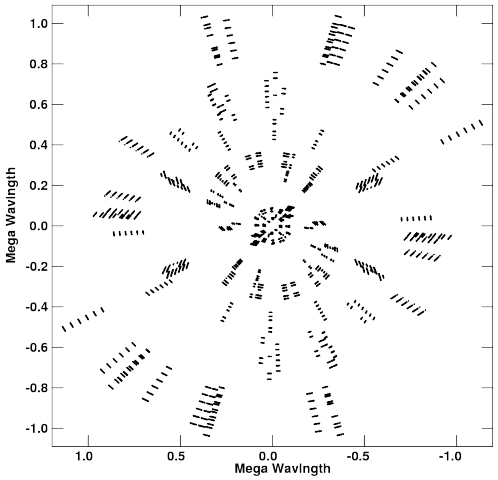
FIGURE B.1: The UV coverages of the seven COBRaS L-band Pointings for the observations taken on Apr. 26th (left) and Apr. 11th (right).



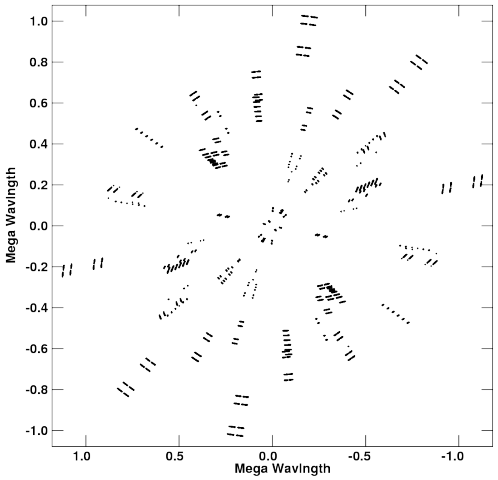

(G) **Point D** ($26^{th}$)



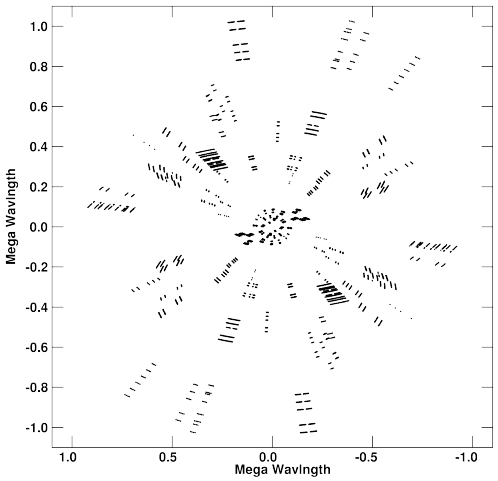

(H) **Point D** ($11^{th}$)



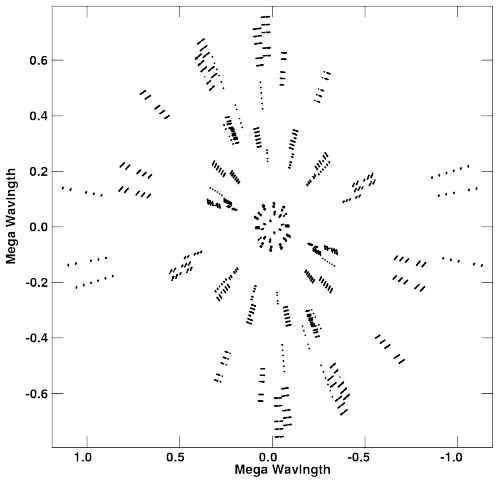

(I) **Point E** ($26^{th}$)




(J) **Point E** ($11^{th}$)

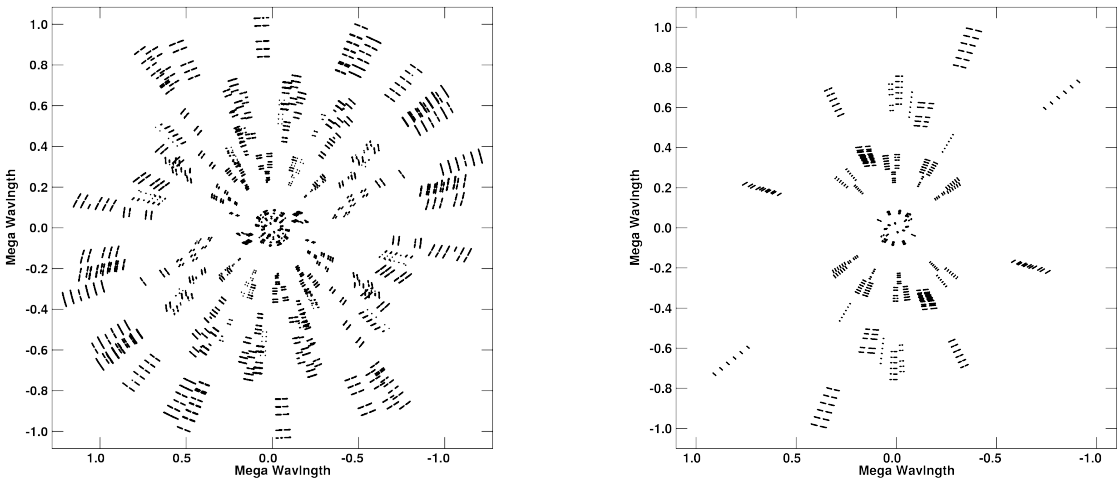



(K) **Point F** ($26^{th}$)




(L) **Point F** ($11^{th}$)

FIGURE B.1: The UV coverages of the seven COBRaS L-band Pointings for the observations taken on Apr. 26th (left) and Apr. 11th (right).




(M) **Point G** ($26^{th}$) (N) **Point G** ($11^{th}$)

FIGURE B.1: The UV coverages of the seven COBRaS L-band Pointings for the observations taken on Apr. 26th (left) and Apr. 11th (right).

# Appendix C

# Additional information regarding the CLASC

Table C.1 found below contains the J2000 positional information (in degrees) for each of the 41 individual sources (47 individual components) found within the CLASC. Their measured flux densities across both COBRaS epochs can be found in Table 3.3. The positional uncertainty associated with each source in RA, $\sigma_{\alpha}$, and DEC, $\sigma_{\delta}$, is also given in Table C.1 in degrees. This was calculated from a combination of the uncertainty in the weighted mean pixel position of each detected island within SEAC, $\sigma_{wmpp}$, and the positional uncertainty in RA and DEC inherent to the observations, $\sigma_{\alpha obs}$ and $\sigma_{\delta obs}$. These in turn, are a function of the signal-to-noise ratio of the detected sources (and therefore the local RMS and peak pixel flux) and the beam size, $\theta^{\alpha}_{beam}$ and $\theta^{\alpha}_{beam}$.

$$\sigma_{\alpha obs} = \frac{\theta^{\alpha}_{beam}}{2SNR} \qquad \text{and} \qquad \sigma_{\delta obs} = \frac{\theta^{\delta}_{beam}}{2SNR} \tag{C.1}$$

where $SNR = F^{peak}_{21cm}/RMS$, i.e. the peak pixel flux divided by the local RMS around the source. Furthermore, Table C.1 contains two columns (one for either observation epoch), which represents the ratio between the number of pixels contributing to its measured flux density (i.e. the source size, $Ss$) and the number of pixels covered by the beam. This gives an indication into how compact ($Ss/\theta_{beam} \sim 1.0 - 1.9$) or extended ($Ss/\theta_{beam} \gtrsim 1.9$) each source is.

TABLE C.1: Positional information and their associated uncertainties of the 41 objects within the COBRaS L-band All Source Catalogue. The last two columns represent the ratio between the number of pixels contributing to its measured flux density (i.e. the source size, $Ss$) and the number of pixels covered by the beam, for the two observation epochs.

| ID | RA (J2000) | $\sigma_\alpha$ | DEC (J2000) | $\sigma_\delta$ | $Ss/\theta_{beam}$ | $Ss/\theta_{beam}$ |
|---|---|---|---|---|---|---|
| | (°) | (°) | (°) | (°) | $26^{th}$ | $11^{th}$ |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 1 | 308.0255439 | 1.3E-06 | 41.2354319 | 1.3E-06 | 4.86 | 3.01 |
| 2 | 308.0345281 | 3.6E-06 | 41.2198489 | 2.9E-06 | 1.91 | ‡0.74 |
| 3 | 308.0592801 | 3.4E-06 | 41.2772928 | 2.8E-06 | 2.10 | ‡0.92 |
| 4 | 308.064438 | 4.9E-06 | 41.2253731 | 3.9E-06 | 1.32 | †1.0 |
| 5 | 308.085808 | 7.9E-06 | 41.005691 | 6.2E-06 | 1.02 | - |
| 6a (Cyg 5) | 308.0934025 | 2.8E-06 | 41.3052367 | 2.3E-06 | 2.57 | - |
| 6b (Cyg 5) | 308.0936687 | 1.5E-06 | 41.3053811 | 1.3E-06 | 5.21 | 1.17 |
| 7 | 308.1119692 | 1.3E-06 | 41.0758238 | 1.2E-06 | 5.89 | 3.44 |
| 8 (A11) | 308.131366 | 3.5E-06 | 41.2355911 | 2.9E-06 | 1.60 | - |
| 9 | 308.1425342 | 2.2E-06 | 41.4039431 | 1.9E-06 | 2.64 | - |
| 10 (SBHW 81) | 308.1527063 | 1.0E-06 | 41.2464921 | 1.0E-06 | 8.91 | 4.52 |
| 11a (SBHW 83) | 308.1590073 | 6.6E-07 | 41.3937504 | 6.0E-07 | 22.28 | 4.29 |
| 11b (SBHW 83) | 308.1590515 | 2.2E-06 | 41.3940838 | 1.8E-06 | 7.00 | - |
| 11c (SBHW 83) | 308.1591699 | 1.4E-06 | 41.394306 | 1.2E-06 | 10.72 | - |
| 12 (Cyg 12) | 308.1706567 | 1.3E-06 | 41.24145 | 1.2E-06 | 5.21 | 2.31 |
| 13 | 308.175348 | 8.6E-06 | 41.4000651 | 6.8E-06 | †1.00 | - |
| 14 | 308.2014459 | 8.2E-06 | 41.051111 | 6.4E-06 | †1.06 | - |
| 15 | 308.2051217 | 7.6E-06 | 41.2471119 | 5.9E-06 | †1.27 | - |
| 16 | 308.208033 | 8.9E-06 | 41.2467119 | 7.0E-06 | ‡0.92 | - |
| 17 (SBHW 90) | 308.2366559 | 6.6E-07 | 41.1481168 | 6.6E-07 | 32.00 | 8.64 |
| 18 | 308.2461174 | 1.8E-06 | 41.0829569 | 1.7E-06 | 3.29 | 1.43 |
| 19 | 308.2819414 | 3.3E-06 | 41.1484881 | 2.7E-06 | 1.89 | - |
| 20 | 308.2899894 | 1.9E-06 | 41.0849299 | 1.7E-06 | 3.06 | ‡0.89 |
| 21 (Cyg 9) | 308.2947166 | 2.9E-06 | 41.2522666 | 2.9E-06 | 3.33 | 1.93 |
| 22 (Cyg 8A) | 308.3127871 | 3.3E-06 | 41.3140058 | 3.3E-06 | 2.78 | 3.09 |
| 23 | 308.3261968 | 4.6E-06 | 41.0365249 | 3.8E-06 | 1.47 | ‡0.77 |
| 24a | 308.3480004 | 6.5E-06 | 41.4570686 | 5.1E-06 | 1.69 | - |
| 24b | 308.3482228 | 1.3E-06 | 41.457102 | 1.2E-06 | 6.93 | - |
| 25 | 308.3482709 | 4.0E-06 | 41.1548469 | 3.3E-06 | 1.61 | - |
| 26 | 308.3624707 | 1.4E-06 | 41.1481299 | 1.4E-06 | 5.00 | 3.32 |
| 27 | 308.3836959 | 2.9E-06 | 41.0993438 | 2.5E-06 | 2.08 | 1.11 |
| 28a | 308.3907761 | 4.4E-06 | 40.9951777 | 3.5E-06 | 1.95 | 1.38 |
| 28b | 308.3911736 | 1.5E-06 | 40.9953888 | 1.3E-06 | 5.84 | - |
| 29 | 308.4022038 | 9.9E-06 | 41.359322 | 7.7E-06 | ‡0.80 | - |
| 30 | 308.4361849 | 1.7E-06 | 41.096148 | 1.5E-06 | 4.39 | 1.06 |
| 31 | 308.4473638 | 5.9E-06 | 41.082443 | 4.7E-06 | 1.34 | - |
| 32 | 308.4666038 | 7.4E-07 | 41.364354 | 6.9E-07 | 14.44 | 2.12 |
| 33a (SBHW 109) | 308.4678032 | 6.4E-06 | 41.2630058 | 5.0E-06 | 1.31 | - |
| 33b (SBHW 109) | 308.4684245 | 3.0E-06 | 41.261717 | 2.4E-06 | 2.87 | - |
| 34 (SBHW 110) | 308.4811667 | 1.5E-06 | 41.0481749 | 1.4E-06 | 5.02 | 4.03 |
| 35 (SBHW 112) | 308.4930229 | 1.2E-06 | 41.1540358 | 1.2E-06 | 5.62 | 3.65 |
| 36 | 308.506144 | 1.7E-06 | 41.0616561 | 1.5E-06 | 3.83 | 1.94 |
| 37 | 308.5101038 | 1.4E-06 | 41.4159412 | 1.3E-06 | 4.98 | 1.90 |
| 38 | 308.5281419 | 2.3E-06 | 41.2668161 | 2.0E-06 | 2.69 | ‡0.90 |
| 39 | 308.5304149 | 7.1E-06 | 41.123505 | 5.6E-06 | 1.09 | - |
| 40 | 308.5755427 | 9.6E-07 | 41.2624131 | 9.0E-07 | 8.65 | - |
| 41 | 308.6363359 | 1.3E-06 | 41.1880419 | 1.2E-06 | 5.75 | - |

† these sources are faint detections and their flux density has been measured with SEAC using threshold values, $\sigma_s = 4.0$ and $\sigma_f = 2.6$.

‡ these sources are very faint detections meaning not only does the same apply to them as those marked with a †, but the detected pixel area is smaller than the beam size, i.e. $Ss/\theta_{beam} < 1.0$.

# Bibliography

Abbott, D. C. et al. (1980). "VLA radio continuum measurements of mass loss from early-type stars". In: *ApJ* 238, pp. 196–202. DOI: 10.1086/157973.

Abbott, David C (1978). "The terminal velocities of stellar winds from early-type stars". In: *ApJ* 225, pp. 893–901.

– (1982). "The theory of radiatively driven stellar winds. II-The line acceleration". In: *ApJ* 259, pp. 282–301.

Abbott, David C., John H. Beiging, and Ed Churchwell (1981). "Mass loss from very luminous OB stars and the Cygnus superbubble". In: *ApJ* 250, pp. 645–659. DOI: 10.1086/159412.

Abbott, David C and LB Lucy (1985). "Multiline transfer and the dynamics of stellar winds". In: *ApJ* 288, pp. 679–693.

Abbott, DC, JH Bieging, and E Churchwell (1984). "The detection of variable, nonthermal radio emission from two O type stars". In: *ApJ* 280, pp. 671–678.

Abbott, DC et al. (1986). "Radio emission from galactic Wolf-Rayet stars and the structure of Wolf-Rayet winds". In: *ApJ* 303, pp. 239–261.

Aihara, H. et al. (2011). "The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III". In: *ApJS* 193, 29, p. 29. DOI: 10.1088/0067-0049/193/2/29. arXiv: 1101.1559 [astro-ph.IM].

Allen, Clabon Walter (1973). *Astrophys. Quant.* Athlone Press, London.

Ambartsumian, V. A. (1954). "On the Origin of Stars". In: *Liege International Astrophysical Colloquia*. Vol. 5. Liege International Astrophysical Colloquia, p. 293.

Argo, Megan (2014). "e-MERLIN data reduction pipeline". In: *Astrophys. Source Code Libr.* 1, p. 07017.

Asplund, Martin et al. (2009). "The chemical composition of the Sun". In: *ARA&A* 47, pp. 481–522.

Athreya, Ramana (2009). "A new approach to mitigation of radio frequency interference in interferometric data". In: *ApJ* 696.1, p. 885.

Baan, WA, PA Fridman, and RP Millenaar (2004). "Radio frequency interference mitigation at the westerbork synthesis radio telescope: Algorithms, test observations, and system implementation". In: *AJ* 128.2, p. 933.

Bally, John, Nickolas Moeckel, and Henry Throop (2005). "Evolution of UV-irradiated protoplanetary disks". In: *Chondrites and the protoplanetary disk*. Vol. 341, p. 81.

Banerjee, S. and P. Kroupa (2015). "Formation of Very Young Massive Clusters and implications for globular clusters". In: *ArXiv e-prints*. arXiv: 1512.03074.

Barkana, Rennan and Abraham Loeb (2001). "In the beginning: the first sources of light and the reionization of the universe". In: *Physics Reports* 349.2, pp. 125–238.

Bate, Matthew R (2000). "Predicting the properties of binary stellar systems: the evolution of accreting protobinary systems". In: *MNRAS* 314.1, pp. 33–53.

Bates, B and S Gilheany (1990). "IUE observations of mass-loss spectral features in B5-B9 supergiants". In: *MNRAS* 243, pp. 320–329.

Beerer, I. M. et al. (2010). "A Spitzer View of Star Formation in the Cygnus X North Complex". In: *ApJ* 720, pp. 679–693. DOI: 10.1088/0004-637X/720/1/679. arXiv: 1007.2112 [astro-ph.SR].

Bell, AR (1978). "The acceleration of cosmic rays in shock fronts". In: *MNRAS* 182.2, pp. 147–156.

Benz, AO and M Güdel (1994). "X-ray/microwave ratio of flares and coronae". In: *A&A* 285.

Beswick, Rob (2013). *e-MERLIN Technical Capabilities Webpage*. URL: http://www.e-merlin.ac.uk/tech/.

Bieging, JH, David C Abbott, and EB Churchwell (1989). "A survey of radio emission from Galactic OB stars". In: *ApJ* 340, pp. 518–536.

Blaauw, Adriaan (1964). "The O associations in the solar neighborhood". In: *ARA&A* 2, p. 213.

Blomme, R. et al. (2005). "Non-thermal radio emission from O-type stars. I. HD168112". In: *A&A* 436, pp. 1033–1040. DOI: 10.1051/0004-6361:20042383. eprint: astro-ph/0502456.

Blomme, R et al. (2010). "Non-thermal radio emission from O-type stars-IV. Cygnus OB2 No. 8A". In: *A&A* 519, A111.

Blomme, R et al. (2013). "The 2.35 year itch of Cygnus OB2# 9-II. Radio monitoring". In: *A&A* 550, A90.

Blomme, Ronny (2011). "Radio observations of massive stars". In: *arXiv preprint arXiv:1102.0386*.

Blomme, Ronny et al. (2003). "Radio and submillimeter observations of wind structure in zeta Pup". In: *A&A* 408, p. 715. DOI: 10.1051/0004-6361:20031007. arXiv: astro-ph/0306310 [astro-ph].

Bohannan, B and PS Conti (1976). "Spectroscopic studies of O-type binaries. I. BD+ 40 4220: an enigma ripe for resolution." In: *ApJ* 204, pp. 797–803.

Bonnell, I. A., S. G. Vine, and M. R. Bate (2004). "Massive star formation: nurture, not nature". In: *MNRAS* 349, pp. 735–741. DOI: 10.1111/j.1365-2966.2004.07543.x. eprint: astro-ph/0401059.

Bouy, H et al. (2013). "Dynamical analysis of nearby clusters-Automated astrometry from the ground: precision proper motions over a wide field". In: *A&A* 554, A101.

Bressert, E et al. (2010). "The spatial distribution of star formation in the solar neighbourhood: do all stars form in dense clusters?" In: *MNRAS* 409.1, pp. L54–L58.

Briggs, FH, JF Bell, and MJ Kesteven (2000). "Removing radio interference from contaminated astronomical spectra using an independent reference signal and closure relations". In: *AJ* 120.6, p. 3351.

Brown, AGA, GdZPT Dekker, and PT De Zeeuw (1997). "Kinematic ages of OB associations". In: *MNRAS* 285.3, pp. 479–492.

Brown, Anthony GA et al. (1999). "OB associations". In: *The Origin of Stars and Planetary Systems*. Springer, pp. 411–440.

Burbidge, GR (1956). "On Synchrotron Radiation from Messier 87." In: *ApJ* 124, p. 416.

Caballero-Nieves, S. M. et al. (2011). "Cyg OB2 Unveiled: The Search for Astrometric Companions". In: *Bulletin de la Societe Royale des Sciences de Liege* 80, pp. 639–643.

Caballero-Nieves, Saida M et al. (2014). "A high angular resolution survey of massive stars in Cygnus OB2: Results from the Hubble space telescope fine guidance sensors". In: *AJ* 147.2, p. 40.

Cash, W. et al. (1980). "The X-ray superbubble in Cygnus". In: *ApJL* 238, pp. L71–L76. DOI: 10.1086/183261.

Cassinelli, Joseph P (1979). "Stellar winds". In: *ARA&A* 17, pp. 275–308.

Castor, John I, David C Abbott, and Richard I Klein (1975). "Radiation-driven winds in Of stars". In: *ApJ* 195, pp. 157–174.

Cazorla, Constantin, Yaël Nazé, and Gregor Rauw (2014). "Wind collisions in three massive stars of Cygnus OB2". In: *A&A* 561, A92.

Chiosi, Cesare and Andre Maeder (1986). "The evolution of massive stars with mass loss". In: *ARA&A* 24, pp. 329–375.

Clark, BG (1980). "An efficient implementation of the algorithm'CLEAN'". In: *A&A* 89, p. 377.

Clark, J. S. et al. (2012). "On the nature of the galactic early-B hypergiants". In: *A&A* 541, p. 145. DOI: 10.1051/0004-6361/201117472.

Clark, JS, VM Larionov, and A Arkharov (2005). "On the population of galactic Luminous Blue Variables". In: *A&A* 435.1, pp. 239–246.

Cohen, David H et al. (2010). "A mass-loss rate determination for $\zeta$ Puppis from the quantitative analysis of X-ray emission-line profiles". In: *MNRAS* 405.4, pp. 2391–2405.

Cohen, David H et al. (2011). "Chandra X-ray spectroscopy of the very early O supergiant HD 93129A: constraints on wind shocks and the mass-loss rate". In: *MNRAS* 415.4, pp. 3354–3364.

Cohen, David H et al. (2014). "Measuring mass-loss rates and constraining shock physics using X-ray line profiles of O stars from the Chandra archive". In: *MNRAS* 439, pp. 908–923.

Colombo, JF Albacete et al. (2007a). "Unveiling the Cygnus OB2 stellar population with Chandra". In: *A&A* 464.1, pp. 211–227.

– (2007b). "Unveiling the Cygnus OB2 stellar population with Chandra". In: *A&A* 464.1, pp. 211–227.

Colombo, JF Albacete et al. (2007c). "X-ray flaring from the young stars in Cygnus OB2". In: *A&A* 474.2, pp. 495–504.

Comerón, F. and A. Pasquali (2012). "New members of the massive stellar population in Cygnus". In: *A&A* 543, A101, A101. DOI: 10.1051/0004-6361/201219022.

Comerón, F. and J. Torra (2001). "Near-infrared imaging of compact HII regions in Cygnus X". In: *A&A* 375, pp. 539–552. DOI: 10.1051/0004-6361:20010654.

Comerón, F et al. (2002). "On the massive star contents of Cygnus OB2". In: *A&A* 389.3, pp. 874–888.

Condon, J. J. et al. (1998). "The NRAO VLA Sky Survey". In: *AJ* 115, pp. 1693–1716. DOI: 10.1086/300337.

Condon, JJ and KJ Mitchell (1984). "A deeper VLA survey of the alpha= 08h52m15s, delta=+ 17 deg 16 arcmin field". In: *AJ* 89, pp. 610–617.

Conti, Peter S and Stewart A Frost (1974). "Variations of the Emission Line Profiles in the O6ef Star Lambda Cephei". In: *ApJ* 190, p. L137.

Contreras, M. E. et al. (1997). "Hipparcos, VLA, and CCD Observations of Cygnus OB2 No. 5: Solving the Mysteryof the Radio "Companion"". In: *ApJL* 488, pp. L153–L156. DOI: 10.1086/310928.

Cornwell, Tim, Robert Braun, and Daniel S Briggs (1999). "Deconvolution". In: *Synthesis Imaging in Radio Astronomy II*. Vol. 180, p. 151.

Cotton, WD (1995). "Fringe fitting". In: *Very Long Baseline Interferometry and the VLBA*. Vol. 82, p. 189.

Cranmer, Steven R and Stanley P Owocki (1996). "Hydrodynamical simulations of corotating interaction regions and discrete absorption components in rotating O-star winds". In: *ApJ* 462, p. 469.

Crowther, Paul (2012). "Birth, life and death of massive stars". In: *Astronomy & Geophysics* 53.4, pp. 4–30.

Crowther, Paul A, Daniel J Lennon, and Nolan R Walborn (2006). "Physical parameters and wind properties of galactic early B supergiants". In: *A&A* 446.1, pp. 279–293.

Crowther, Paul A et al. (2000). "Far Ultraviolet Spectroscopic Explorer spectroscopy of the O VI resonance doublet in Sand 2 (WO)". In: *ApJL* 538.1, p. L51.

Crowther, Paul A et al. (2016). "The R136 star cluster dissected with Hubble Space Telescope/STIS. I. Far-ultraviolet spectroscopic census and the origin of He ii $\lambda$1640 in young star clusters". In: *MNRAS*, stw273.

Cutri, R. M. et al. (2003). "VizieR Online Data Catalog: 2MASS All-Sky Catalog of Point Sources (Cutri+ 2003)". In: *VizieR Online Data Catalog* 2246.

Dale, J. E., B. Ercolano, and I. A. Bonnell (2012). "Ionization-induced star formation - IV. Triggering in bound clusters". In: *MNRAS* 427, pp. 2852–2865. DOI: 10.1111/j.1365-2966.2012.22104.x. arXiv: 1208.4486.

Dame, T. M. et al. (1987). "A composite CO survey of the entire Milky Way". In: *ApJ* 322, pp. 706–720. DOI: 10.1086/165766.

Dame, Thomas M, Dap Hartmann, and P Thaddeus (2001). "The Milky Way in molecular clouds: A new complete CO survey". In: *ApJ* 547.2, p. 792.

Dame, TM and P Thaddeus (1985). "A wide-latitude CO survey of molecular clouds in the northern Milky Way". In: *ApJ* 297, pp. 751–765.

De Becker, Michaël (2007). "Non-thermal emission processes in massive binaries". In: *A&AR* 14.3-4, pp. 171–216.

De Becker, Michaël and Françoise Raucq (2013). "Catalogue of particle-accelerating colliding-wind binaries". In: *A&A* 558, A28.

De Becker, Michaël, Grégor Rauw, and Jean Manfroid (2004). "A Spectroscopic study of the non-thermal radio emitter Cyg OB2# 8A: discovery of a new binary system". In: *A&A* 424.3, pp. L39–L42.

De Becker, Michaël et al. (2006). "XMM–Newton observations of the massive colliding wind binary and non-thermal radio emitter Cyg OB2# 8A [O6If+ O5. 5III (f)]". In: *MNRAS* 371.3, pp. 1280–1294.

De Mink, SE, OR Pols, and RW Hilditch (2007). "Efficiency of mass transfer in massive close binaries-Tests from double-lined eclipsing binaries in the SMC". In: *A&A* 467.3, pp. 1181–1196.

De Ruiter, HR, AG Willis, and HC Arp (1977). "A Westerbork 1415 MHz survey of background radio sources. II-Optical identifications with deep IIIA-J plates". In: *A&ASS* 28, pp. 211–293.

Dessart, Luc and SP Owocki (2005). "2D simulations of the line-driven instability in hot-star winds-II. Approximations for the 2D radiation force". In: *A&A* 437.2, pp. 657–666.

Dickel, H. R. and H. J. Wendker (1978). "The Cygnus X region. XI - Map of visual extinction A sub V". In: *A&A* 66, pp. 289–295.

Dixon, Robert S (1970). "A master list of radio sources". In: *ApJSS* 20, pp. 1–503.

Dougherty, S. M. and P. M. Williams (2000). "Non-thermal emission in Wolf-Rayet stars: are massive companions required?" In: *MNRAS* 319, pp. 1005–1010. DOI: 10.1046/j.1365-8711.2000.03837.x.

Dougherty, Sean M and Julian M Pittard (2006). "Winds in collision: high-energy particles in massive binary systems". In: *arXiv preprint astro-ph/0611088*.

Drew, J. E. et al. (2005). "The INT Photometric H$\alpha$ Survey of the Northern Galactic Plane (IPHAS)". In: *MNRAS* 362, pp. 753–776. DOI: 10.1111/j.1365-2966.2005.09330.x. eprint: astro-ph/0506726.

Drew, J. E. et al. (2008). "Early-A stars from IPHAS, and their distribution in and around the Cyg OB2 association". In: *MNRAS* 386, pp. 1761–1773. DOI: 10.1111/j.1365-2966.2008.13147.x. arXiv: 0802.3868.

Drew, JE (1989). "Ionization and thermal equilibrium models for O star winds based on time-independent radiation-driven wind theory". In: *ApJSS* 71, pp. 267–291.

– (1990). "A new theoretical calibration of the relation between mass-loss rate and H-alpha emission for O star winds". In: *ApJ* 357, pp. 573–581.

Dzib, Sergio A et al. (2013). "Multi-epoch Very Long Baseline Array Observations of the Compact Wind-collision Region in the Quadruple System Cyg OB2# 5". In: *ApJ* 763.2, p. 139.

Ebbets, D (1980). "The variable H-alpha P Cygni profile of Alpha Camelopardalis-O9. 5 IA". In: *ApJ* 235, pp. 97–103.

Eggleton, P. (2006). *Evolutionary Processes in Binary and Multiple Stars*.

Eggleton, Peter P (1983). "Approximations to the radii of Roche lobes". In: *ApJ* 268, p. 368.

Ekström, S. et al. (2012). "Grids of stellar models with rotation. I. Models from 0.8 to 120 M at solar metallicity (Z = 0.014)". In: *A&A* 537, A146, A146. DOI: 10.1051/0004-6361/201117751. arXiv: 1110.5049 [astro-ph.SR].

Elder, FR et al. (1947). "Radiation from electrons in a synchrotron". In: *Physical Review* 71.11, p. 829.

Elmegreen, B. G. (2007). "On the Rapid Collapse and Evolution of Molecular Clouds". In: *ApJ* 668, pp. 1064–1082. DOI: 10.1086/521327. arXiv: 0707.2252.

Elmegreen, B. G. and D. A. Hunter (2000). "A Pressure Anomaly for H II Regions in Irregular Galaxies". In: *ApJ* 540, pp. 814–824. DOI: 10.1086/309382. eprint: astro-ph/0005182.

Elmegreen, Bruce G and Deidre A Hunter (2010). "On the Disruption of Star Clusters in a Hierarchical Interstellar Medium". In: *ApJ* 712.1, p. 604.

Evans, CJ et al. (2004). "The ultraviolet and optical spectra of luminous B-type stars in the Small Magellanic Cloud". In: *PASP* 116.824, p. 909.

Eversberg, Thomas, Anthony FJ Moffat, et al. (1998). "Outmoving Clumps in the Wind of the Hot O Supergiant? Puppis". In: *ApJ* 494.2, p. 799.

Fall, S. M., R. Chandar, and B. C. Whitmore (2005). "The Age Distribution of Massive Star Clusters in the Antennae Galaxies". In: *ApJL* 631, pp. L133–L136. DOI: 10.1086/496878. eprint: astro-ph/0509293.

Farnier, C, R Walter, and J-C Leyder (2011). "$\eta$Carinae: a very large hadron collider". In: *A&A* 526, A57.

Feldmeier, A (1995). "Time-dependent structure and energy transfer in hot star winds." In: *A&A* 299, p. 523.

Feldmeier, A, J Puls, and AWA Pauldrach (1997). "A possible origin for X-rays from O stars." In: *A&A* 322, pp. 878–895.

Feldmeier, Achim and Isaac Shlosman (2000). "Runaway of line-driven winds toward critical and overloaded solutions". In: *ApJL* 532.2, p. L125.

Fitzgerald, M. P. (1970). "The Intrinsic Colours of Stars and Two-Colour Reddening Lines". In: *A&A* 4, p. 234.

Fitzpatrick, E. L. and D. Massa (2007). "An Analysis of the Shapes of Interstellar Extinction Curves. V. The IR-through-UV Curve Morphology". In: *ApJ* 663, pp. 320–341. DOI: 10.1086/518158. arXiv: 0705.0154.

Fridman, PA and WA Baan (2001). "RFI mitigation methods in radio astronomy". In: *A&A* 378.1, pp. 327–344.

Friend, David B and David C Abbott (1986). "The theory of radiatively driven stellar winds. III-Wind models with finite disk correction and rotation". In: *ApJ* 311, pp. 701–707.

Friend, David B and JI Castor (1983). "Stellar winds driven by multiline scattering". In: *ApJ* 272, pp. 259–272.

Fullerton, Alexander W., D. L. Massa, and R. K. Prinja (2006). "The discordance of mass-loss estimates for galactic o-type stars". In: *ApJ* 637, pp. 1025–1039. DOI: 10.1086/498560. arXiv: astro-ph/0510252 [astro-ph].

Fullerton, AW et al. (1997). "Wind variability of B supergiants. III. Corotating spiral structures in the stellar wind of HD 64760." In: *A&A* 327, pp. 699–720.

Galván-Madrid, R. et al. (2013). "MUSCLE W49: A Multi-Scale Continuum and Line Exploration of the Most Luminous Star Formation Region in the Milky Way. I. Data and the Mass Structure of the Giant Molecular Cloud". In: *ApJ* 779, 121, p. 121. DOI: 10.1088/0004-637X/779/2/121. arXiv: 1309.4129.

Garay, Guido et al. (2003). "A triple radio continuum source associated with IRAS 16547-4247: A collimated stellar wind emanating from a massive protostar". In: *ApJ* 587.2, p. 739.

Gayley, Kenneth G (1995). "An improved line-strength parameterization in hot-star winds". In: *ApJ* 454, p. 410.

Geller, Aaron M et al. (2015). "Different Dynamical Ages for the Two Young and Coeval LMC Star Clusters, NGC 1805 and NGC 1818, Imprinted on Their Binary Populations". In: *ApJ* 805.1, p. 11.

Gieles, Mark and Simon F Portegies Zwart (2011). "The distinction between star clusters and associations". In: *Monthly Notices of the Royal Astronomical Society: Letters* 410.1, pp. L6–L7.

Goodwin, Simon P and Nate Bastian (2006). "Gas expulsion and the destruction of massive young clusters". In: *MNRAS* 373.2, pp. 752–758.

Gottschalk, M. et al. (2012). "The JCMT $^{12}$CO(3-2) survey of the Cygnus X region. I. A pathfinder". In: *A&A* 541, A79, A79. DOI: 10.1051/0004-6361/201118600. arXiv: 1202.0832.

Grunhut, J H et al. (2016). "The MiMeS survey of magnetism in massive stars: Magnetic analysis of the O-type stars". In: *The MiMeS survey of magnetism in massive stars: Magnetic analysis of the O-type stars*. eprint arXiv:1610.07895.

Guarcello, M. G. et al. (2012). "Optical Photometric GTC/OSIRIS Observations of the Young Massive Association Cygnus OB2". In: *ApJS* 202, 19, p. 19. DOI: 10.1088/0067-0049/202/2/19. arXiv: 1208.6296 [astro-ph.SR].

Guarcello, M. G. et al. (2015). "Optical and infrared counterparts of the X-ray sources detected in the Chandra Cygnus OB2 Legacy Survey". In: *ArXiv e-prints*. arXiv: 1501.03761 [astro-ph.SR].

Guarcello, MG et al. (2013). "The Protoplanetary Disks in the Nearby Massive Star-forming Region Cygnus OB2". In: *AJ* 773.2, p. 135.

Güdel, M. et al. (2002). "Detection of the Neupert Effect in the Corona of an RS Canum Venaticorum Binary System by XMM-Newton and the Very Large Array". In: *ApJ* 577, pp. 371–376. DOI: 10.1086/342122. eprint: astro-ph/0206080.

Gürkan, M. A., M. Freitag, and F. A. Rasio (2004). "Formation of Massive Black Holes in Dense Star Clusters. I. Mass Segregation and Core Collapse". In: *ApJ* 604, pp. 632–652. DOI: 10.1086/381968. eprint: astro-ph/0308449.

Gutermuth, R. A. et al. (2009). "A Spitzer Survey of Young Stellar Clusters within One Kiloparsec of the Sun: Cluster Core Extraction and Basic Structural Analysis". In: *ApJS* 184, pp. 18–83. DOI: 10.1088/0067-0049/184/1/18. arXiv: 0906.0201 [astro-ph.SR].

Gvaramadze, VV and DJ Bomans (2008). "Search for OB stars running away from young star clusters-I. NGC 6611". In: *A&A* 490.3, pp. 1071–1077.

Haehnelt, Martin G et al. (2001). "An ionizing ultraviolet background dominated by massive stars". In: *ApJL* 549.2, p. L151.

Hales, Christopher A et al. (2012). "BLOBCAT: software to catalogue flood-filled blobs in radio images of total intensity and linear polarization". In: *MNRAS* 425.2, pp. 979–996.

Hall, DS (1974). "The O9+ Of eclipsing binary V729 Cygni in Cygnus OB2". In: *Acta Astronomica* 24, p. 69.

Hancock, PJ et al. (2012). "Aegean: Compact source finding in radio images". In: *Astrophysics Source Code Library* 1, p. 12009.

Hanson, M. M. (2003). "A Study of Cygnus OB2: Pointing the Way toward Finding Our Galaxy's Super-Star Clusters". In: *ApJ* 597, pp. 957–969. DOI: 10.1086/378508. eprint: astro-ph/0307540.

Hanson, MM (2003). "A Study of Cygnus OB2: Pointing the Way toward Finding Our Galaxy's Super-Star Clusters". In: *ApJ* 597.2, p. 957.

Harnden Jr., F. R. et al. (1979). "Discovery of an X-ray star association in VI Cygni /Cyg OB2/". In: *ApJL* 234, pp. L51–L54. DOI: 10.1086/183107.

Haser, Stefan M et al. (1998). "Quantitative UV spectroscopy of early O stars in the Magellanic Clouds. The determination of the stellar metallicities". In: *A&A* 330, pp. 285–305.

Heger, A, N Langer, and SE Woosley (2000). "Presupernova evolution of rotating massive stars. I. Numerical method and evolution of the internal stellar structure". In: *ApJ* 528.1, p. 368.

Heger, A, SE Woosley, and HC Spruit (2005). "Presupernova evolution of differentially rotating massive stars including magnetic fields". In: *ApJ* 626.1, p. 350.

Heggie, D. C. (1975). "Binary evolution in stellar dynamics". In: *MNRAS* 173, pp. 729–787. DOI: 10.1093/mnras/173.3.729.

Henrichs, HF, L Kaper, and GAA Zwarthoed (1988). "Rapid variability in O star winds". In: *A Decade of UV Astronomy with the IUE Satellite* 2, pp. 145–149.

Herrero, A., F. Najarro, and J. Puls (2003). "Spectroscopic Analysis of Cyg OB2 Supergiants Using Unified Model Atmospheres". In: *Stellar Atmosphere Modeling*. Vol. 288. ASP Conf. Ser. P. 243.

Herrero, A, J Puls, and F Najarro (2002). "Fundamental parameters of Galactic luminous OB stars VI. Temperatures, masses and WLR of Cyg OB2 supergiants". In: *A&A* 396.3, pp. 949–966.

Herrero, A., J. Puls, and M. R. Villamariz (2000). "Fundamental parameters of Galactic luminous OB stars. IV. The upper HR diagram". In: *A&A* 354, pp. 193–215. eprint: astro-ph/9912200.

Herrero, A et al. (2001). "An analysis of HST UV spectra of Cyg OB2 stars". In: *A&A* 366.2, pp. 623–635.

Hewett, P. C. et al. (2006). "The UKIRT Infrared Deep Sky Survey ZY JHK photometric system: passbands and synthetic colours". In: *MNRAS* 367, pp. 454–468. DOI: 10.1111/j.1365-2966.2005.09969.x. eprint: astro-ph/0601592.

Hillier, D John and Daniel L Miller (1998). "The treatment of non-LTE line blanketing in spherically expanding outflows". In: *ApJ* 496.1, p. 407.

Hills, J. G. (1975). "Encounters between binary and single stars and their effect on the dynamical evolution of stellar systems". In: *AJ* 80, pp. 809–825. DOI: 10.1086/111815.

Hills, JG (1980). "The effect of mass loss on the dynamical evolution of a stellar system-Analytic approximations". In: *ApJ* 235, pp. 986–991.

Högbom, JA (1974). "Aperture synthesis with a non-regular distribution of interferometer baselines". In: *A&ASS* 15, p. 417.

Hopkins, A Mobasher et al. (1998). "The PHOENIX deep survey: 1.4-GHz source counts". In: *MNRAS* 296.4, pp. 839–846.

Howarth, Ian D and Raman K Prinja (1989). "The stellar winds of 203 Galactic O stars-A quantitative ultraviolet survey". In: *ApJSS* 69, pp. 527–592.

Humphreys, R. M. (1978). "Studies of luminous stars in nearby galaxies. I. Supergiants and O stars in the Milky Way." In: *ApJS* 38, pp. 309–350. DOI: 10.1086/190559.

Humphreys, Roberta M and Kris Davidson (1994). "The luminous blue variables: Astrophysical geysers". In: *PASP*, pp. 1025–1051.

Hutchings, J. B. (1981). "Stellar winds in Cygnus OB2 /VI Cygni/". In: *PASP* 93, pp. 50–54. DOI: 10.1086/130774.

Ignace, Richard (2016). "Long-wavelength, free–free spectral energy distributions from porous stellar winds". In: *MNRAS* 457.4, pp. 4123–4134.

Johnson, H. L. and W. W. Morgan (1954). "A Heavily Obscured O-Association in Cygnus." In: *ApJ* 119, p. 344. DOI: 10.1086/145831.

Kaper, L et al. (1999). "Long-and short-term variability in O-star winds II. Quantitative analysis of DAC behaviour". In: *A&A* 344, pp. 231–262.

Keenan, Philip Childs (1971). "Classification of supergiants of types G, K, and M." In: *Contr. from the Kitt Peak Natl. Obser.* 554, pp. 35–39.

Kennedy, M et al. (2010). "Modeling the Radio Emission from Cyg OB2 No. 5: A Quadruple System?" In: *ApJ* 709.2, p. 632.

Kennicutt Jr, Robert C (1998). "Star formation in galaxies along the Hubble sequence". In: *ARA&A* 36, p. 189.

Kiminki, Daniel C and Henry A Kobulnicky (2012). "An updated look at binary characteristics of massive stars in the Cygnus OB2 association". In: *ApJ* 751.1, p. 4.

Kiminki, Daniel C, M Virginia McSwain, and Henry A Kobulnicky (2008). "New Massive Binaries in the Cygnus OB2 Association". In: *ApJ* 679.2, p. 1478.

Kiminki, Daniel C et al. (2007). "A radial velocity survey of the Cyg OB2 association". In: *ApJ* 664.2, p. 1102.

Kiminki, Daniel C et al. (2009). "Five More Massive Binaries in the Cygnus OB2 Association". In: *AJ* 137.6, p. 4608.

Kiminki, Daniel C et al. (2012). "Additional Massive Binaries in the Cygnus OB2 Association". In: *ApJ* 747.1, p. 41.

Kiminki, Daniel C et al. (2015). "PREDICTING GAIA? S PARALLAX DISTANCE TO THE CYGNUS OB2 ASSOCIATION WITH ECLIPSING BINARIES". In: *ApJ* 811.2, p. 85.

Kitamoto, Shunji and Koji Mukai (1996). "ASCA observation of Cygnus OB2 association". In: *Publications of the Astronomical Society of Japan* 48.6, pp. 813–818.

Knödlseder, J (2000). "Cygnus OB2-a young globular cluster in the Milky Way". In: *A&A* 360, pp. 539–548.

Kobulnicky, H. A. et al. (2014). "Toward Complete Statistics of Massive Binary Stars: Penultimate Results from the Cygnus OB2 Radial Velocity Survey". In: *ApJS* 213, 34, p. 34. DOI: 10.1088/0067-0049/213/2/34. arXiv: 1406.6655 [astro-ph.SR].

Kobulnicky, Henry A, Ian J Gilbert, and Daniel C Kiminki (2010). "OB Stars and Stellar Bow shocks in Cygnus-X: A Novel Laboratory Estimating Stellar Mass Loss Rates". In: *ApJ* 710.1, p. 549.

Kobulnicky, Henry A et al. (2012). "A Fresh Catch of Massive Binaries in the Cygnus OB2 Association". In: *ApJ* 756.1, p. 50.

Koornneef, J. (1983). "Near-infrared photometry. II - Intrinsic colours and the absolute calibration from one to five micron". In: *A&A* 128, pp. 84–93.

Kouwenhoven, MBN and R De Grijs (2008). "The effect of binaries on the dynamical mass determination of star clusters". In: *A&A* 480.1, pp. 103–114.

Kroupa, P. (2001). "On the variation of the initial mass function". In: *MNRAS* 322, pp. 231–246. DOI: 10.1046/j.1365-8711.2001.04022.x. eprint: astro-ph/0009005.

– (2008). "Initial Conditions for Star Clusters". In: *The Cambridge N-Body Lectures*. Ed. by S. J. Aarseth, C. A. Tout, and R. A. Mardling. Vol. 760. Lecture Notes in Physics, Berlin Springer Verlag, p. 181. DOI: 10.1007/978-1-4020-8431-7_8. arXiv: 0803.1833.

Krtička, Jiří and Jiri Kubat (2006). "The winds of hot massive first stars". In: *A&A* 446.3, pp. 1039–1049.

Kruijssen, J. M. D. et al. (2011). "Modelling the formation and evolution of star cluster populations in galaxy simulations". In: *MNRAS* 414, pp. 1339–1364. DOI: 10.1111/j.1365-2966.2011.18467.x. arXiv: 1102.1013.

Kruijssen, J. M. D. et al. (2012). "Formation versus destruction: the evolution of the star cluster population in galaxy mergers". In: *MNRAS* 421, pp. 1927–1941. DOI: 10.1111/j.1365-2966.2012.20322.x. arXiv: 1112.1065.

Krumholz, M. R., R. I. Klein, and C. F. McKee (2012). "Radiation-hydrodynamic Simulations of the Formation of Orion-like Star Clusters. II. The Initial Mass Function from Winds, Turbulence, and Radiation". In: *ApJ* 754, 71, p. 71. DOI: 10.1088/0004-637X/754/1/71. arXiv: 1203.2620 [astro-ph.SR].

Kudritzki, Rolf P (2002). "Line-driven winds, ionizing fluxes, and ultraviolet spectra of hot stars at extremely low metallicity. I. Very massive O stars". In: *ApJ* 577.1, p. 389.

Kudritzki, Rolf-Peter and Joachim Puls (2000). "Winds from hot stars". In: *ARA&A* 38.1, pp. 613–666.

Kudritzki, RP et al. (1989). "Radiation-driven winds of hot stars. VI-Analytical solutions for wind models including the finite cone angle effect". In: *A&A* 219, pp. 205–218.

Kudritzki, RP et al. (1991). "Winds of Hot Stars in the Magellanic Clouds". In: *The Magellanic Clouds*. Vol. 148. Springer Science & Business Media, p. 279.

Kuhn, M. A. et al. (2014). "The Spatial Structure of Young Stellar Clusters. I. Subclusters". In: *ApJ* 787, 107, p. 107. DOI: 10.1088/0004-637X/787/2/107. arXiv: 1403.4252.

Kühr, H et al. (1979). "A catalogue of radio sources". In: *Preprint* 55.

Kurapati, Sushma et al. (2016). "A JVLA survey of the high frequency radio emission of the massive magnetic B-and O-type stars". In: *MNRAS*, stw2838.

Lada, C. J. and E. A. Lada (2003). "Embedded Clusters in Molecular Clouds". In: *ARA&A* 41, pp. 57–115. DOI: 10.1146/annurev.astro.41.011802.094844. eprint: astro-ph/0301540.

Lada, CJ, M Margulis, and D Dearborn (1984). "The formation and early dynamical evolution of bound stellar systems". In: *ApJ* 285, pp. 141–152.

Lamers, Henny JGLM and Claus Leitherer (1993). "What are the mass-loss rates of O stars?" In: *ApJ* 412, pp. 771–791.

Lamers, HJGLM, M Cerruti-Sola, and M Perinotto (1987). "The'SEI'method for accurate and efficient calculations of line profiles in spherically symmetric stellar winds". In: *ApJ* 314, pp. 726–738.

Larson, Richard B (1969). "Numerical calculations of the dynamics of a collapsing proto-star". In: *MNRAS* 145.3, pp. 271–295.

Lawrence, L. C. and V. C. Reddish (1965). "The Cygnus II association. I. Intercomparison of photometries with the Edinburgh and Hamburg Schmidt telescopes." In: *Publications of the Royal Observatory of Edinburgh* 3, pp. 275–308.

Leitherer, C. and C. Robert (1991). "Observations of stellar winds from hot stars at 1.3 millimeters". In: *ApJ* 377, pp. 629–638. DOI: 10.1086/170390.

Leitherer, C. et al. (1982). "Spectroscopy and infrared photometry of CYG OB 2 stars - Velocity law and mass-loss rates". In: *A&A* 108, pp. 102–110.

Leitherer, Claus, Carmelle Robert, and Laurent Drissen (1992). "Deposition of mass, momentum, and energy by massive stars into the interstellar medium". In: *ApJ* 401, pp. 596–617.

Lemmon, John J (1997). "Wideband model of man-made HF noise and interference". In: *Radio Sci* 32.2, pp. 525–539.

Leung, K-C and DP Schneider (1978). "Evolved contact systems of spectral type O. III-V729 Cygni". In: *ApJ* 224, pp. 565–569.

Linder, Natacha et al. (2009). "A multiwavelength investigation of the massive eclipsing binary Cygnus OB2# 5". In: *A&A* 495.1, pp. 231–241.

Ling, Zhixing, Shuang Nan Zhang, and Shichao Tang (2009). "Determining the distance of Cyg X-3 with its X-ray dust scattering halo". In: *ApJ* 695.2, p. 1111.

Lis, D. C. et al. (1994). "Star formation in the galactic center dust ridge". In: *ApJ* 423, pp. L39–L42. DOI: 10.1086/187230.

Lobel, A and R Blomme (2008). "Modeling ultraviolet wind line variability in massive hot stars". In: *ApJ* 678.1, p. 408.

Longmore, S. N. et al. (2012). "G0.253 + 0.016: A Molecular Cloud Progenitor of an Arches-like Cluster". In: *ApJ* 746, 117, p. 117. DOI: 10.1088/0004-637X/746/2/117. arXiv: 1111.3199.

Longmore, Steven N et al. (2014). "The formation and early evolution of young massive clusters". In: *Protostars and Planets VI* 1, pp. 291–314.

Lucas, P. W. et al. (2008). "The UKIDSS Galactic Plane Survey". In: *MNRAS* 391, pp. 136–163. DOI: 10.1111/j.1365-2966.2008.13924.x. arXiv: 0712.0100.

Lucy, L B_ and RL White (1980). "X-ray emission from the winds of hot stars". In: *ApJ* 241, pp. 300–305.

Lucy, LB and PM Solomon (1970). "Mass loss by hot stars". In: *ApJ* 159, p. 879.

MacGregor, KB, Lo Hartmann, and JC Raymond (1979). "Radiative amplification of sound waves in the winds of O and B stars". In: *ApJ* 231, pp. 514–523.

Maeder, André and Georges Meynet (1987). "Grids of evolutionary models of massive stars with mass loss and overshooting-Properties of Wolf-Rayet stars sensitive to overshooting". In: *A&A* 182, pp. 243–263.

– (2010). "Evolution of massive stars with mass loss and rotation". In: *New Astronomy Reviews* 54.3, pp. 32–38.

Mahy, L. et al. (2013). "A spectroscopic investigation of the O-type star population in four Cygnus OB associations. I. Determination of the binary fraction". In: *A&A* 550, A27, A27. DOI: 10.1051/0004-6361/201219425. arXiv: 1301.0500 [astro-ph.SR].

Maia, F. F. S., E. Moraux, and I. Joncour (2016). "Young and embedded clusters in Cygnus-X: evidence for building up the initial mass function?" In: *MNRAS* 458, pp. 3027–3046. DOI: 10.1093/mnras/stw450. arXiv: 1602.07341 [astro-ph.SR].

Maíz-Apellániz, Jesús et al. (2004). "A Galactic O star catalog". In: *ApJSS* 151.1, p. 103.

Marcote, Benito et al. (2015). "Physical properties of the gamma-ray binary LS 5039 through low-and high-frequency radio observations". In: *MNRAS* 451.1, pp. 4578–4592.

Markova, N and J Puls (2008). "Bright OB stars in the Galaxy-IV. Stellar and wind parameters of early to late B supergiants". In: *A&A* 478.3, pp. 823–842.

Martí, J, JM Paredes, and M Peracaula (2001). "Development of a two-sided relativistic jet in Cygnus X-3". In: *A&A* 375.2, pp. 476–484.

Martí, J et al. (2005). "Possible hot spots excited by the relativistic jets of Cygnus X-3". In: *A&A* 439.1, pp. 279–285.

Martí, J. et al. (2007). "Deep radio images of the HEGRA and Whipple TeV sources in the Cygnus OB2 region". In: *A&A* 472, pp. 557–564. DOI: 10.1051/0004-6361:20077712. arXiv: 0705.3801.

Martins, F. and B. Plez (2006). "UBVJHK synthetic photometry of Galactic O stars". In: *A&A* 457, pp. 637–644. DOI: 10.1051/0004-6361:20065753. eprint: astro-ph/0606587.

Martins, F., D. Schaerer, and D. J. Hillier (2005). "A new calibration of stellar parameters of Galactic O stars". In: *A&A* 436.3, p. 17. DOI: 10.1051/0004-6361:20042386.

Martins, Fabrice, Daniel Schaerer, and D John Hillier (2002). "On the effective temperature scale of O stars". In: *A&A* 382.3, pp. 999–1004.

Martins, Fabrice et al. (2005). "O stars with weak winds: the Galactic case". In: *A&A* 441.2, pp. 735–762.

Maryeva, O. V. et al. (2016). "On the nature of high reddening of Cygnus OB2 #12 hypergiant". In: *MNRAS* 458, pp. 491–507. arXiv: 1602.05042 [astro-ph.SR].

Maryeva, Olga, VG Klochkova, and EL Chentsov (2013). "Modeling of atmospheres of the brightest stars belonging to the Cyg OB2 association". In: *Massive Stars: From alpha to Omega*, p. 84.

Maryeva, OV and R Ya Zhuchkov (2012). "Medium resolution spectroscopy of the supergiant O31f Cyg OB2 No. 7". In: *Astrophys.* 55.3, pp. 371–376.

Masias, M et al. (2012). "A review of source detection approaches in astronomical images". In: *MNRAS* 422.2, pp. 1674–1689.

Mason, Brian D et al. (1998). "ICCD speckle observations of binary stars. XIX. An astrometric/spectroscopic survey of O stars". In: *AJ* 115.2, p. 821.

Massa, Derck et al. (1995). "The IUE Mega Campaign: Wind variability and rotation in early-type stars". In: *ApJL* 452.1, p. L53.

Massey, Philip and AB Thompson (1991). "Massive stars in CYG OB2". In: *AJ* 101, pp. 1408–1428.

McKee, C. F. and J. C. Tan (2003). "The Formation of Massive Stars from Turbulent Cores". In: *ApJ* 585, pp. 850–871. DOI: 10.1086/346149. eprint: astro-ph/0206037.

Menten, KM, T Pillai, and F Wyrowski (2005). "Initial conditions for massive star birth–Infrared dark clouds". In: *Proceedings of the International Astronomical Union* 1.S227, pp. 23–34.

Meurer, G. R. et al. (1995). "Starbursts and Star Clusters in the Ultraviolet". In: *AJ* 110, p. 2665. DOI: $10.1086/117721$. eprint: astro-ph/9509038.

Michelson, Albert A (1890). "I. On the application of interference methods to astronomical measurements". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 30.182, pp. 1–21.

Miralles, M Paz et al. (1994). "VLA observations of Cyg OB2 No. 5: detection of a weak radio companion and new observations and models for the main component". In: *A&A* 282, pp. 547–553.

Mokiem, M. R. et al. (2005). "Spectral analysis of early-type stars using a genetic algorithm based fitting method "". In: *A&A* 441.2, p. 23. DOI: 10.1051/0004-6361:20053522.

Mokiem, MR et al. (2007). "The empirical metallicity dependence of the mass-loss rate of O-and early B-type stars". In: *A&A* 473.2, pp. 603–614.

Morford, JC et al. (2016). "e-MERLIN 21 cm constraints on the mass-loss rates of OB stars in Cyg OB2". In: *MNRAS* 463.1, pp. 763–770.

Morgan, William Wilson, Philip Childs Keenan, and Edith Kellman (1943). "An atlas of stellar spectra, with an outline of spectral classification". In: *Chicago, Ill., The University of Chicago press [1943]* 1.

Morgan, WW, AD Code, and AE Whitford (1955). "Studies in Galactic STRUCTURE. II. LUMINOSITY Classification for 1270 Blue Giant Stars." In: *ApJSS* 2, p. 41.

Morton, Donald C (1967a). "Mass loss from three OB supergiants in Orion". In: *ApJ* 150, p. 535.

– (1967b). "The far-ultraviolet spectra of six stars in Orion". In: *ApJ* 147, p. 1017.

Motte, Frédérique et al. (2005). "The earliest phases of massive star formation within entire molecular cloud complexes". In: *Proceedings of the International Astronomical Union* 1.S227, pp. 151–156.

Muijres, LE et al. (2011). "Predictions of the effect of clumping on the wind properties of O-type stars". In: *A&A* 526, A32.

Mullan, DJ (1984). "Asymmetries in stellar MG II H and K and CA II H and K line profiles-Discrepancies between MG and CA asymmetries". In: *ApJ* 284, pp. 769–773.

Münch, L. and W. W. Morgan (1953). "Notes: A Probable Clustering of Blue Giants in Cygnus." In: *ApJ* 118, pp. 161–162. DOI: 10.1086/145737.

Murray, N. (2011). "Star Formation Efficiencies and Lifetimes of Giant Molecular Clouds in the Milky Way". In: *ApJ* 729, 133, p. 133. DOI: 10.1088/0004-637X/729/2/133. arXiv: 1007.3270.

Najarro, F, MM Hanson, and J Puls (2011). "L-band spectroscopy of Galactic OB-stars". In: *A&A* 535, A32.

Najarro, F et al. (2008). "Tracking the Clumping in OB Stars from UV to radio". In: *Clumping in Hot-Star Winds*. Vol. 1, p. 43.

Nazé, Yaël et al. (2008). "A binary signature in the non-thermal radio-emitter Cyg OB2# 9". In: *A&A* 483.2, pp. 543–546.

Nazé, Yaël et al. (2010). "FIRST ORBITAL SOLUTION FOR THE NON-THERMAL EMITTER Cyg OB2 NO. 9Based on observations collected at the Haute-Provence Observatory and with XMM-Newton, an ESA Science Mission with instruments and contributions directly funded by ESA Member States and the USA (NASA)." In: *ApJ* 719.1, p. 634.

Nazé, Yaël et al. (2012). "The 2.35 year itch of Cygnus OB2# 9-I. Optical and X-ray monitoring". In: *A&A* 546, A37.

Negueruela, Ignacio et al. (2008). "New very massive stars in Cygnus OB2". In: *A&A* 487.2, pp. 575–581.

Nguyen-Lu'o'ng, Q. et al. (2013). "Low-velocity Shocks Traced by Extended SiO Emission along the W43 Ridges: Witnessing the Formation of Young Massive Clusters". In: *ApJ* 775, 88, p. 88. DOI: 10.1088/0004-637X/775/2/88. arXiv: 1306.0547 [astro-ph.GA].

Noerdlinger, Peter D (2008). "Solar mass loss, the astronomical unit, and the scale of the solar system". In: *arXiv preprint arXiv:0801.3807*.

Nugis, T, PA Crowther, and AJ Willis (1998). "Clumping-corrected mass-loss rates of Wolf-Rayet stars". In: *A&A* 333, pp. 956–969.

Nugis, T and HJGLM Lamers (2000). "Mass-loss rates of Wolf-Rayet stars as a function of stellar parameters". In: *A&A* 360, pp. 227–244.

– (2002). "The mass-loss rates of Wolf–Rayet stars explained by optically thick radiation driven wind models". In: *A&A* 389.1, pp. 162–179.

Offner, S. S. R., C. E. Hansen, and M. R. Krumholz (2009). "Stellar Kinematics of Young Clusters in Turbulent Hydrodynamic Simulations". In: *ApJ* 704, pp. L124–L128. DOI: 10.1088/0004-637X/704/2/L124. arXiv: 0909.4304 [astro-ph.SR].

Offringa, AR et al. (2010). "Post-correlation radio frequency interference classification methods". In: *MNRAS* 405.1, pp. 155–167.

Olnon, FM (1975). "Thermal bremsstrahlung radiospectra for inhomogeneous objects, with an application to MWC 349." In: *A&A* 39, pp. 217–223.

Oskinova, LM, W-R Hamann, and A Feldmeier (2007). "Neglecting the porosity of hot-star winds can lead to underestimating mass-loss rates". In: *A&A* 476.3, pp. 1331–1340.

Osterbrock, Donald and Edith Flather (1959). "Electron Densities in the Orion NEBULA. II." In: *ApJ* 129, p. 26.

Owocki, SP, SR Cranmer, and KG Gayley (1996). "Inhibition of wind-compressed disk formation by nonradial line forces in rotating hot-star winds". In: *ApJL* 472.2, p. L115.

Owocki, SP and GB Rybicki (1984). "Instabilities in line-driven stellar winds. I-Dependence on perturbation wavelength". In: *ApJ* 284, pp. 337–350.

– (1985). "Instabilities in line-driven stellar winds. II-Effect of scattering". In: *ApJ* 299, pp. 265–276.

Owocki, Stanley (2014). "Theory of Winds from Hot, Luminous Massive Stars". In: *arXiv preprint arXiv:1409.2084*.

Owocki, Stanley P, John I Castor, and George B Rybicki (1988). "Time-dependent models of radiatively driven stellar winds. I-Nonlinear evolution of instabilities for a pure absorption model". In: *ApJ* 335, pp. 914–930.

Owocki, Stanley P and David H Cohen (2001). "X-ray line profiles from parameterized emission within an accelerating stellar wind". In: *ApJ* 559.2, p. 1108.

Owocki, Stanley P, Steven R Cranmer, and Alexander W Fullerton (1995). "Periodic variations in ultraviolet spectral lines of the B0. 5 Ib Star HD 64760: evidence for corotating wind streams rooted in surface variations". In: *ApJL* 453.1, p. L37.

Owocki, Stanley P, Kenneth G Gayley, and Nir J Shaviv (2004). "A porosity-length formalism for photon-tiring-limited mass loss from stars above the Eddington limit". In: *ApJ* 616.1, p. 525.

Owocki, Stanley P, Richard HD Townsend, et al. (2008). "Dynamical Simulations of Magnetically Channeled Line-Driven Stellar Winds: III. Angular Momentum Loss and Rotational Spindown". In: *arXiv preprint arXiv:0810.4247*.

Paczynski, B (1971). "Evolutionary processes in close binary systems". In: *ARA&A* 9, p. 183.

Palacio, Santiago del et al. (2016). "A model for the non-thermal emission of the very massive colliding-wind binary HD 93129A". In: *arXiv preprint arXiv:1604.07294*.

Panagia, Nino and Marcello Felli (1975). "The spectrum of the free-free radiation from extended envelopes". In: *A&A* 39, pp. 1–5.

Parker, R. J. and S. P. Goodwin (2007). "Do O-stars form in isolation?" In: *MNRAS* 380, pp. 1271–1275. DOI: 10.1111/j.1365-2966.2007.12179.x. arXiv: 0707.0605.

Parkin, ER et al. (2014). "The 2.35 year itch of Cygnus OB2# 9-III. X-ray and radio emission analysis based on 3D hydrodynamical modelling". In: *A&A* 570, A10.

Pauldrach, A (1987). "Radiation driven winds of hot luminous stars. III-Detailed statistical equilibrium calculations for hydrogen to zinc". In: *A&A* 183, pp. 295–313.

Pauldrach, A, TL Hoffmann, and M Lennon (2001). "Radiation-driven winds of hot luminous stars-XIII. A description of NLTE line blocking and blanketing towards realistic models for expanding atmospheres". In: *A&A* 375.1, pp. 161–195.

Pauldrach, A, J Puls, and RPz Kudritzki (1986). "Radiation-driven winds of hot luminous stars-Improvements of the theory and first results". In: *A&A* 164, pp. 86–100.

Pauldrach, AWA et al. (1994). "Radiation-driven winds of hot luminous stars. 12: A first step towards detailed UV-line diagnostics of O-stars". In: *A&A* 283, pp. 525–560.

Peck, Luke W (2014). "Technical Developmet and Scientific Preparation for the e-MERLIN Cygnus OB2 Radio Survey". PhD thesis. University College London.

Peck, Luke W and Danielle M Fenech (2013). "SERPent: Automated reduction and RFI-mitigation software for e-MERLIN". In: *Astron. & Comput.* 2, pp. 54–66.

Pelupessy, F. I. and S. Portegies Zwart (2012). "The evolution of embedded star clusters". In: *MNRAS* 420, pp. 1503–1517. DOI: 10.1111/j.1365-2966.2011.20137.x. arXiv: 1111.0992 [astro-ph.SR].

Perley, Rick A and Bryan J Butler (2013). "An accurate flux density scale from 1 to 50 GHz". In: *ApJSS* 204.2, p. 19.

Perryman, MAC et al. (2001). "GAIA: Composition, formation and evolution of the Galaxy". In: *A&A* 369.1, pp. 339–363.

Persi, P. et al. (1985). "Time-variable, excess radio emission from CYG OB2 No. 5". In: *A&A* 142, pp. 263–267.

Petit, V et al. (2013). "A magnetic confinement versus rotation classification of massive-star magnetospheres". In: *MNRAS* 429.398, pp. 1211–0282.

Petrov, Blagovest, Jorick S Vink, and Götz Gräfener (2016). "Two bi-stability jumps in theoretical wind models for massive stars and the implications for luminous blue variable supernovae". In: *MNRAS* 458.2, pp. 1999–2011.

Pflamm-Altenburg, Jan and Pavel Kroupa (2010). "The two-step ejection of massive stars and the issue of their formation in isolation". In: *MNRAS* 404.3, pp. 1564–1568.

Piddington, J. H. and H. C. Minnett (1952). "Radio-Frequency Radiation from the Constellation of Cygnus". In: *Australian Journal of Scientific Research A Physical Sciences* 5, p. 17.

Pigulski, A. and Z. Kołaczkowski (1998). "Mapping the $\beta$ Cephei instability strip: photometric variability of stars in the central part of the Cygnus OB2 association". In: *MNRAS* 298, pp. 753–762. DOI: 10.1046/j.1365-8711.1998.01650.x.

Pittard, JM and SM Dougherty (2006). "Radio, X-ray, and $\gamma$-ray emission models of the colliding-wind binary WR 140". In: *MNRAS* 372.2, pp. 801–826.

Pittard, Julian M and Ian R Stevens (1997). "Theoretical X-ray properties of colliding stellar winds in O+ O star binaries". In: *MNRAS* 292.2, pp. 298–316.

Plume, René et al. (1997). "Dense gas and star formation: Characteristics of cloud cores associated with water masers". In: *ApJ* 476.2, p. 730.

Podsiadlowski, Ph, PC Joss, and JJL Hsu (1992). "Presupernova evolution in massive interacting binaries". In: *ApJ* 391, pp. 246–264.

Podsiadlowski, Philipp (2001). "Common-envelope evolution and stellar mergers". In: *Evolution of Binary and Multiple Star Systems*. Vol. 229, p. 239.

Pols, OR and M Marinus (1994). "Monte-Carlo simulations of binary stellar evolution in young open clusters". In: *A&A* 288.

Portegies Zwart, S. F. and S. L. W. McMillan (2002). "The Runaway Growth of Intermediate-Mass Black Holes in Dense Star Clusters". In: *ApJ* 576, pp. 899–907. DOI: 10.1086/341798. eprint: astro-ph/0201055.

Portegies Zwart, S. F., S. L. W. McMillan, and M. Gieles (2010). "Young Massive Star Clusters". In: *ARA&A* 48, pp. 431–493. DOI: 10.1146/annurev-astro-081309-130834. arXiv: 1002.1961.

Price, D. J. and M. R. Bate (2009). "Inefficient star formation: the combined effects of magnetic fields and radiative feedback". In: *MNRAS* 398, pp. 33–46. DOI: 10.1111/j.1365-2966.2009.14969.x. arXiv: 0904.4071 [astro-ph.SR].

Prinja, R. K., D. Massa, and S. C. Searle (2005). "The ionization structure of early-B supergiant winds". In: *A&A* 430, pp. L41–L44. DOI: 10.1051/0004-6361:200400129. eprint: astro-ph/0412315.

Prinja, Raman K (1988). "Evidence for rotationally modulated variability in O star winds". In: *MNRAS* 231.1, 21P–24P.

– (1994). "Time-Dependent Phenomena in OB Star Winds". In: *Pulsation, Rotation and Mass Loss in Early-Type Stars*. Springer, pp. 507–516.

Prinja, Raman K, MJ Barlow, and Ian D Howarth (1990). "Terminal velocities for a large sample of O stars, B supergiants, and Wolf-Rayet stars". In: *ApJ* 361, pp. 607–620.

Prinja, RK, AW Fullerton, and PA Crowther (1996). "Variability in the optical wind lines of HD 151804 (O8 Iaf)." In: *A&A* 311, pp. 264–272.

Prinja, RK and DL Massa (2010). "Signature of wide-spread clumping in B supergiant winds". In: *A&A* 521, p. L55.

Prinja, RK, DL Massa, and M Cantiello (2012). "LOOKING DEEP INTO THE CAT'S EYE: STRUCTURE AND ROTATION IN THE FAST WIND OF THE PN CENTRAL STAR OF NGC 6543Based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with program 12489." In: *ApJL* 759.2, p. L28.

Prinja, RK and LJ Smith (1992). "Migrating optical depth enhancements in the UV wind lines of the Wolf-Rayet star HD 93131". In: *A&A* 266, pp. 377–384.

Pshirkov, MS, VV Vasiliev, and KA Postnov (2016). "Evidence of Fermi bubbles around M31". In: *MNRASL* 459.1, pp. L76–L80.

Puls, J, U Springmann, and M Lennon (2000). "Radiation driven winds of hot luminous stars-XIV. Line statistics and radiative driving". In: *A&ASS* 141.1, pp. 23–64.

Puls, J et al. (1996). "O-star mass-loss and wind momentum rates in the Galaxy and the Magellanic Clouds Observations and theoretical predictions." In: *A&A* 305, p. 171.

Puls, Joachim, Jorick S Vink, and Francisco Najarro (2008a). "Mass loss from hot massive stars". In: *A&AR* 16.3-4, pp. 209–325.

– (2008b). "Mass loss from hot massive stars". In: *A&AR* 16.3-4, pp. 209–325.

Puls, Joachim et al. (2006). "Bright OB stars in the Galaxy. 3. Constraints on the radial stratification of the clumping factor in hot star winds from a combined Halpha, IR and radio analysis". In: *A&A* 454, pp. 625–651. DOI: 10.1051/0004-6361:20065073. arXiv: astro-ph/0604372 [astro-ph].

Rathborne, JM, JM Jackson, and R Simon (2006). "Infrared dark clouds: precursors to star clusters". In: *ApJ* 641.1, p. 389.

Rauw, G. (2011). "A multi-epoch XMM-Newton campaign on the core of the massive Cygnus OB2 association". In: *A&A* 536, A31, A31. DOI: 10.1051/0004-6361/201117648. arXiv: 1110.4223 [astro-ph.SR].

Rauw, G. et al. (2002). "A multi-wavelength investigation of the non-thermal radio emitting O-star 9 Sgr". In: *A&A* 394, pp. 993–1008. DOI: 10.1051/0004-6361:20020926.

Rauw, Grégor, J-M Vreux, and B Bohannan (1999). "The Interacting Early-Type Binary BD+ 40 ° 4220 (V729 Cyg): Modeling the Colliding Winds RegionAll the observations reported here were obtained at the Observatoire de Haute Provence and at the Kitt Peak National Observatory." In: *ApJ* 517.1, p. 416.

Reddish, V. C., L. C. Lawrence, and N. N. Pratt (1967). "The Cygnus II association-II. The distribution of stars and interstellar matter". In: *MNRAS* 136, pp. 428–429. DOI: 10.1093/mnras/136.4.428.

Reimer, A, M Pohl, and O Reimer (2006). "Nonthermal high-energy emission from colliding winds of massive stars". In: *ApJ* 644.2, p. 1118.

Reipurth, Bo and Nicola Schneider (2008). "Star Formation and Young Clusters in Cygnus". In: *Handbook of Star Forming Regions* 1, p. 36.

Repolust, T et al. (2005). "Quantitative H and K band spectroscopy of Galactic OB-stars at medium resolution". In: *A&A* 440.1, pp. 261–286.

Rios, LY and K DeGioia-Eastwood (2004). "A New Very Early Contact Binary in Cyg OB2". In: *Bulletin of the American Astronomical Society*. Vol. 205, p. 905.

Runacres, MC and SP Owocki (2002). "The outer evolution of instability-generated structure in radiatively driven stellar winds". In: *A&A* 381.3, pp. 1015–1025.

Rygl, KLJ et al. (2012). "Parallaxes and proper motions of interstellar masers toward the Cygnus X star-forming complex-I. Membership of the Cygnus X region". In: *A&A* 539, A79.

Sana, H and CJ Evans (2011). "Active OB Stars: Structure, Evolution, Mass Loss, and Critical Limits". In: *IAU Symp.* Vol. 272.

Sana, Hugues, G James, and Eric Gosset (2011). "The massive star binary fraction in young open clusters–III. IC 2944 and the Cen OB2 association". In: *MNRAS* 416.2, pp. 817–831.

Sana, Hugues et al. (2012). "Binary interaction dominates the evolution of massive stars". In: *Science* 337.6093, pp. 444–446.

Santolaya-Rey, AE, J Puls, and A Herrero (1997). "Atmospheric NLTE-models for the spectroscopic analysis of luminous blue stars with winds." In: *A&A* 323, pp. 488–512.

Schilbach, E and S Röser (2008). "On the origin of field O-type stars". In: *A&A* 489.1, pp. 105–114.

Schneider, N. et al. (2006). "A new view of the Cygnus X region. KOSMA $^{13}$CO 2 to 1, 3 to 2, and $^{12}$CO 3 to 2 imaging". In: *A&A* 458, pp. 855–871. DOI: 10.1051/0004-6361:20065088.

Schneider, N. et al. (2007). "A multiwavelength study of the S106 region. III. The S106 molecular cloud as part of the Cygnus X cloud complex". In: *A&A* 474, pp. 873–882. DOI: 10.1051/0004-6361:20077540.

Schulte, D. H. (1956). "New Members of the Association VI Cygni." In: *ApJ* 124, p. 530. DOI: 10.1086/146256.

– (1958). "New Members of the Association VI Cygni. II." In: *ApJ* 128, p. 41. DOI: 10.1086/146513.

Schwab, FR (1984). "Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry". In: *ApJ* 89, pp. 1076–1081.

Scuderi, S. et al. (1998). "Radio observations of stellar winds from early type stars". In: *A&A* 332, pp. 251–267.

Searle, Samantha C et al. (2008). "Quantitative studies of the optical and UV spectra of Galactic early B supergiants-I. Fundamental parameters". In: *A&A* 481.3, pp. 777–797.

Setia Gunawan, Diah YA et al. (2003). "A Westerbork Synthesis Radio Telescope 1400 and 350 MHz continuum survey of the cygnus OB2 association, in search of Hot Massive Stars". In: *ApJS* 149, pp. 123–156.

Shirley, Yancy L et al. (2003). "A CS J= 5 4 mapping survey toward high-mass star-forming cores associated with water masers". In: *ApJSS* 149.2, p. 375.

Shu, F. H., F. C. Adams, and S. Lizano (1987). "Star formation in molecular clouds - Observation and theory". In: *ARA&A* 25, pp. 23–81. DOI: 10.1146/annurev.aa.25.090187.000323.

Skiff, BA (2009). "VizieR Online Data Catalog: Catalogue of Stellar Spectral Classifications (Skiff, 2010)". In: *VizieR Online Data Catalog* 1, p. 02023.

Sobolev, Viktor Viktorovich (1960). "Moving envelopes of stars". In: *Cambridge: Harvard University Press, 1960* 1.

Sokolov, Arseniĭ Aleksandrovich and Igor Mikhaĭlovich Ternov (1966). "Synchrotron radiation". In: *Akademia Nauk SSSR, Moskovskoie Obshchestvo Ispytatelei prirody. Sektsia Fiziki. Sinkhrotron Radiation, Nauka Eds., Moscow, 1966 (Russian title: Sinkhrotronnoie izluchenie), 228 pp.* 1.

Sonneborn, George and JC Bouret (2011). "Near-Infrared Mass Loss Diagnostics for Massive Stars". In: *Bulletin of the American Astronomical Society*. Vol. 43, p. 15408.

Souza, S. P. and B. L. Lutz (1980). "On the nature of Cygnus OB2 number 12". In: *ApJL* 235, pp. L87–L90. DOI: 10.1086/183164.

Spitzer Jr, Lyman (1958). "Distribution of Galactic Clusters." In: *ApJ* 127, p. 17.

Spitzer Jr, Lyman and Michael H Hart (1971). "Random gravitational encounters and the evolution of spherical systems. I. Method". In: *ApJ* 164, p. 399.

Stevens, Ian R, John M Blondin, and AMT Pollock (1992). "Colliding winds from early-type stars in binary systems". In: *ApJ* 386, pp. 265–287.

Strom, S. E., K. M. Strom, and G. L. Grasdalen (1975). "Young stellar objects and dark interstellar clouds". In: *ARA&A* 13, pp. 187–216. DOI: 10.1146/annurev.aa.13.090175.001155.

Sundqvist, J. O. et al. (2011). "Mass loss from inhomogeneous hot star winds. II. Constraints from a combined optical/UV study". In: *A&A* 528, A64, A64. DOI: 10.1051/0004-6361/201015771. arXiv: 1101.5293 [astro-ph.SR].

Sundqvist, JO, J Puls, and SP Owocki (2014). "Mass loss from inhomogeneous hot star winds-III. An effective-opacity formalism for line radiative transfer in accelerating, clumped two-component media, and first results on theory and diagnostics". In: *A&A* 568, A59.

Šurlan, B et al. (2013). "Macroclumping as solution of the discrepancy between H$\alpha$ and P v mass loss diagnostics for O-type stars". In: *A&A* 559, A130.

Tak, FFS Van der and KM Menten (2005). "Very compact radio emission from high-mass protostars-II. Dust disks and ionized accretion flows". In: *A&A* 437.3, pp. 947–956.

Tavani, M et al. (2009). "Detection of gamma-ray emission from the eta-carinae region". In: *ApJL* 698.2, p. L142.

Taylor, MB (2005). "TOPCAT & STIL: starlink table VOTable processing software". In: *Astronomical Data Analysis Software and Systems XIV*. Vol. 347, p. 29.

– (2006). "STILTS: A Package for Command-Line Processing of Tabular Data". In: *Astronomical Data Analysis Software and Systems XV*. Vol. 351, p. 666.

Thompson, A Richard (1999). "Fundamentals of radio interferometry". In: *Synthesis Imaging in Radio Astronomy II*. Vol. 180, p. 11.

Torres-Dodgen, A. V., M. Carroll, and M. Tapia (1991). "UVBY and JHKL-prime photometry of OB stars in the association Cygnus OB2". In: *MNRAS* 249, pp. 1–12. DOI: 10.1093/mnras/249.1.1.

Townsend, Richard HD, Stanley P Owocki, et al. (2006). "Centrifugal breakout of magnetically confined line-driven stellar winds". In: *ApJL* 640.2, p. L191.

Ud-Doula, Asif, Stanley P Owocki, and Richard HD Townsend (2008). "Dynamical simulations of magnetically channelled line-driven stellar winds–II. The effects of field-aligned rotation". In: *MNRAS* 385.1, pp. 97–108.

Uyanıker, B. et al. (2001). "The Cygnus superbubble revisited". In: *A&A* 371, pp. 675–697. DOI: 10.1051/0004-6361:20010387.

Van Genderen, AM (2001). "S Doradus variables in the Galaxy and the Magellanic Clouds". In: *A&A* 366.2, pp. 508–531.

Van Loo, S, MC Runacres, and R Blomme (2005). "A layered model for non-thermal radio emission from single O stars". In: *A&A* 433.1, pp. 313–322.

van Loo, S. et al. (2008). "Non-thermal radio emission from O-type stars. III. Is Cygnus OB2 No. 9 a wind-colliding binary?" In: *A&A* 483, pp. 585–590. DOI: 10.1051/0004-6361:200809367. arXiv: 0803.2607.

Vink, J. S. et al. (2008). "IPHAS discoveries of young stars towards Cyg OB2 and its southern periphery". In: *MNRAS* 387, pp. 308–318. DOI: 10.1111/j.1365-2966.2008.13220.x. arXiv: 0803.2840.

Vink, Jorick S., A. de Koter, and H. J. Lamers (2001). "Mass-loss predictions for O and B stars as a function of metallicity "". In: *A&A* 369.2, p. 15. DOI: 10.1051/0004-6361:20010127.

Vink, Jorick S, Alex de Koter, and Henny J Lamers (1999). "On the nature of the bi-stability jump in the winds of early-type supergiants". In: *A&A* 350, pp. 181–196.

Vink, Jorick S, Alex de Koter, and Henny JGLM Lamers (2000). "New theoretical mass-loss rates of O and B stars". In: *A&A* 362, pp. 395–309.

Voelcker, K. (1975). "Infrared observations of the association CYG OB 2". In: *A&AS* 22, pp. 1–7.

Vreux, J-M (1985). "Emission-line variability in WR stars: short periods and nonradial pulsations?" In: *PASP* 97.589, p. 274.

Wade, GA et al. (2016). "The MiMeS survey of magnetism in massive stars: introduction and overview". In: *MNRAS* 456.1, pp. 2–22.

Walborn, N. R. (1973). "Spectral Classification of Stars in VI Cygni (cygnus OB2)". In: *ApJL* 180, p. L35. DOI: 10.1086/181147.

Walborn, NR (1980). "An Atlas of yellow-red OB spectra". In: *ApJSS* 44, pp. 535–538.

Waldron, Wayne L et al. (1998). "X-ray and radio observations of the Cygnus OB2 association". In: *ApJSS* 118.1, p. 217.

Waltman, EB et al. (1994). "The quiescent level of Cygnus X-3 at 2.25 and 8.3 GHz: 1988-1992". In: *AJ* 108, pp. 179–187.

Waters, LBFM and HJGLM Lamers (1984). "The infrared and radio spectrum of early type stars with mass loss. II-Tables of theoretical curves of growth for IR and radio excess and gaunt factors". In: *A&AS* 57, pp. 327–352.

Weber, R et al. (1997). "Spectral detector for interference time blanking using quantized correlator". In: *A&ASS* 126.1, pp. 161–167.

Wellstein, S, N Langer, and H Braun (2001). "Formation of contact in massive close binaries". In: *A&A* 369.3, pp. 939–959.

White, Richard L (1985). "Synchrotron emission from chaotic stellar winds". In: *Radio Stars*. Springer, pp. 45–46.

White, RL and RH Becker (1983). "The discovery of a hot stellar wind". In: *ApJ* 272, pp. L19–L23.

Williams, JP, L Blitz, and CF McKee (2000). *Protostars and Planets IV, ed.*

Williams, PM et al. (1990). "Multi-frequency variations of the Wolf-Rayet system HD 193793. I-Infrared, X-ray and radio observations". In: *MNRAS* 243, pp. 662–684.

Wit, W. J. de et al. (2005). "The Origin of massive O-type field stars. 2. Field O stars as runaways". In: *A&A* 437, pp. 247–255. DOI: 10.1051/0004-6361:20042489. arXiv: astro-ph/0503337 [astro-ph].

Wolff, SC et al. (2007). "Rotational Velocities for B0-B3 Stars in Seven Young Clusters: Further Study of the Relationship between Rotation Speed and Density in Star-Forming Regions". In: *AJ* 133.3, p. 1092.

Woosley, Stanford E, Alex Heger, and Thomas A Weaver (2002). "The evolution and explosion of massive stars". In: *Reviews of Modern Physics* 74.4, p. 1015.

Woźniak, PR et al. (2004). "Northern Sky Variability Survey: Public Data ReleaseBased on observations obtained with the ROTSE-I robotic telescope, which was operated at Los Alamos National Laboratory." In: *AJ* 127.4, p. 2436.

Wright, A. E. and M. J. Barlow (1975). "The radio and infrared spectrum of early-type stars undergoing mass loss". In: *MNRAS* 170, pp. 41–51.

Wright, N. J. and J. J. Drake (2009). "The Massive Star-Forming Region Cygnus OB2. I. Chandra Catalog of Association Members". In: *ApJS* 184, pp. 84–99. DOI: 10.1088/0067-0049/184/1/84. arXiv: 0908.0549 [astro-ph.SR].

Wright, N. J. et al. (2014a). "Constraints on massive star formation: Cygnus OB2 was always an association". In: *MNRAS* 438, pp. 639–646. DOI: 10.1093/mnras/stt2232. arXiv: 1311.4537 [astro-ph.SR].

Wright, N. J. et al. (2014b). "The Chandra Cygnus OB2 Legacy Survey: Design and X-ray Point Source Catalog". In: *ArXiv e-prints*. arXiv: 1408.6579 [astro-ph.SR].

Wright, Nicholas J, Janet E Drew, and Michael Mohr-Smith (2015). "The massive star population of Cygnus OB2". In: *MNRAS* 449.1, pp. 741–760.

Wright, Nicholas J et al. (2010). "The massive star-forming region Cygnus OB2. II. Integrated stellar properties and the star formation history". In: *ApJ* 713.2, p. 871.

Wright, Nicholas J et al. (2014). "Constraints on massive star formation: Cygnus OB2 was always an association". In: *MNRAS* 438.1, pp. 639–646.

Wright, Nicholas J et al. (2016). "Cygnus OB2 DANCe: A high precision proper motion study of the Cygnus OB2 association". In: *MNRAS*, stw1148.

Wyithe, J Stuart B and Abraham Loeb (2003). "Was the Universe Reionized by Massive Metal-free Stars?" In: *ApJL* 588.2, p. L69.

Zinnecker, H. and H. W. Yorke (2007). "Toward Understanding Massive Star Formation". In: *ARA&A* 45, pp. 481–563. DOI: 10.1146/annurev.astro.44.051905.092549. arXiv: 0707.1279.